

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 22, 2005, 10:39:15 / Search time 5453 Seconds  
(without alignments)  
5691.642 Million cell updates/sec

Title: US-09-651-651-5  
Sequence: 1 MCANSKSVYATSFYATVAFPL.....DSESIPIKRTAVWELDKSGY 546

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5883141 segs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DRV=xlh  
-Q=/cgn2.1/USFTO.spool/US09651651/runat.18112005.101830.20245/app.query.fasta.1.711  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rig -MINMATCH=0.1 -LCOFCL=0 -LCOFEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=psco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09651651 @CEN 1.1 3367 @runat.18112005.101830.20245 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_hlg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2965	100.0	1641	AX090311 Sequence
2	2944	99.3	1899	AY989885 Arabidops
3	2332	78.7	3896	BD271623 Novel c1a

4	2332	78.7	3896	6	BD271633	BD271633 Novel c1a
5	2332	78.7	3896	6	AX037587	AX037587 Sequence
6	2332	78.7	3896	6	AX037606	AX037606 Sequence
7	2332	78.7	119914	15	AC003027	AC003027 Arabidops
8	2318	78.2	2609	15	AF493159	AF493159 Medicago
9	2300.5	77.6	2531	15	AY957608	AY957608 Citrus si
10	1483.5	50.0	110000	15	AP008208	Continuation (349
11	1483.5	50.0	110000	15	AP004120	Continuation (350
12	1483.5	50.0	118192	15	AP005691	AP005691 Oryza sat
13	1483.5	50.0	147123	15	AP004882	AP004882 Oryza sat
14	1483.5	50.0	158971	14	AP004882	AP004882 Citrus cor
15	915	30.9	93398	15	AP006356	AP006356 Lotus cor
16	755.5	25.5	86556	14	AP008162	AP008162 Lotus cor
17	444	15.0	328	6	AX090328	AX090328 Sequence
18	356	12.0	110000	15	AE016819	AE016819 Ashbya go
19	356	12.0	110000	15	AE016819	Continuation (2 of
20	346	11.7	110000	15	CR954209	Continuation (5 of
21	343	11.6	3877	15	AF293457	AF293458 Zea mays
22	339	11.4	1893	6	CO891270	CO891270 Sequence
23	325.5	11.1	110000	15	CR382126	Continuation (7 of
24	325.5	11.0	110000	15	CR382131	Continuation (20 o
25	324	10.9	2106	6	AR546016	AR546016 Sequence
26	316.5	10.7	110000	15	CR382136	Continuation (8 of
27	316.5	10.7	110000	15	CR382136	Continuation (9 of
28	299.5	10.1	1986	6	BD271616	BD271616 Novel c1a
29	299.5	10.1	1986	6	BD271625	BD271625 Novel c1a
30	299.5	10.1	1986	6	BD271626	BD271626 Novel c1a
31	299.5	10.1	1986	6	AX037577	AX037577 Sequence
32	299.5	10.1	1986	6	AX037595	AX037595 Sequence
33	299.5	10.1	1986	6	AX037597	AX037597 Sequence
34	299.5	10.1	1986	6	AX090382	AX090382 Sequence
35	299.5	10.1	2575	15	SCYNR008M	Z71623 S.cerevisia
36	299.5	10.1	23901	15	SCN201952	X77395 S.cerevisia
37	292	9.8	110000	15	CNS07EGJ	AL590462 DNA cent
38	290.5	9.8	103568	15	CNS07EGJ	AX090329 Sequence
39	290.5	9.8	322194	15	CNS09S4S	AX090329 Sequence
40	288	9.7	356	6	AX090329	AX090329 Sequence
41	284	9.6	2398	6	AX794712	AX794712 Sequence
42	283.5	9.6	2700	6	AX794710	AX794710 Sequence
43	281	9.5	2030	6	AX794714	AX794714 Sequence
44	280	9.4	2479	6	AX794704	AX794704 Sequence
45	278	9.4	2312	6	BD271617	BD271617 Novel c1a

## ALIGNMENTS

RESULT 1	AX090311	1641 bp	DNA	linear	PAT 21-MAR-2001
LOCUS	AX090311				
DEFINITION	Sequence 4 from Patent WO0116308.				
ACCESSION	AX090311				
VERSION	AX090311.1	GI:13444180			
KEYWORDS					
SOURCE					
ORGANISM	Arabidopsis thaliana (thale cress)				
REFERENCE	1				
AUTHORS	Lassner, M. and van Eenennaam, A.				
TITLE	Plant sterol acyltransferases				
JOURNAL	Patent: WO 0116308-A 4 08-MAR-2001;				
FEATURES	MONSANTO COMPANY (US)				
source	Location/Qualifiers				
	1..1641				
	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				

ORIGIN	Alignment Scores:	5.65e-246	Length:	1641
	Pred. No.:	2965.00	Matches:	546
	Score:			

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-651-651-5 (1-546) x AX090311 (1-1641)

QY 1 MetGlyAlaAsnSerIysSerValThrAlaSerPheThrValIleAlaValaPhePheLeu 20  
 Db 1 AAGGAGCGCAATTCGAATTCAGTAACGGCTTCCTTCACCGTCATCGCCCTTTTTCCTG 60  
 QY 21 IleCysGlyGlyArgThrAlaValaGluAspGluThrGluPheHisGlyAspTyrSerIys 40  
 Db 61 ATTTCGGGTGGCCGAACTCGCGGTGAGAGATGAGACCGAATTCACGGCCACTACTCGAAG 120  
 QY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSerIleLeu 60  
 Db 121 CTATCGGGATATATCATTCCTGGGATTTGGCGTCGAGCAGCTAGAGCGTGGTCAATCCTT 180  
 QY 61 AspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrThr 80  
 Db 181 GACTGCCATACACTCCGTTGACCTTCAATCCGCTCGACCTCGTATGCTAGACACACT 240  
 QY 81 LysLeuLeuSerAlaValaLysCysTrpPheLysCysMetValLeuAspProTyrAsnGln 100  
 Db 241 AAGCTCTCTTCTGCTGCTCACTGCTGCTGTTAACTGATGCTAGATCTCTATATCA 300  
 QY 101 ThrAspHisProGluCysAlaSerSerArgProAspSerGlyLeuSerAlaIleThrGluLeu 120  
 Db 301 ACAGACCATCCGAGAGTAAAGTCACCGCGCTGACAGGTGCTTTTCAGCCATCACAGAAATTG 360  
 QY 121 AspProGlyTyrIleThrGlyProLeuSerThrValTrpLysGluTrpLeuLysTrpCys 140  
 Db 361 GATCCAGGTTTACATAACAGGTCCTCTTCTACTGCTGTGAAGAAGGCGTTAAGTGGT 420  
 QY 141 ValGluPheGlyIleGluIleAsnAlaIleValAlaValProTyrAspTrpArgLeuSer 160  
 Db 421 GTTAGATTGGTATAGAACCAATGCAATTCGCTGCTCCATACGATTTGGAGATTGTA 480  
 QY 161 ProThrLysLeuGluGluArgAspLeuTyrPheHisLysLeuLysLeuThrPheGluThr 180  
 Db 481 CCACCAAAATTGGAAAGAGCGTGACCTTTACTTTCACCAAGCTCAAGTTGACCTTTGAAACT 540  
 QY 181 AlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHisSerMetCysAsnAsnVal 200  
 Db 541 GCTTTAAAACCTCCGTGGCGGCCCTTCTATAGATTTGCCCATTCATGGGTAATATATGTC 600  
 QY 201 PheArgTyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisTrpTyrLeuLysTrp 220  
 Db 601 TTCAGATACTTTCTGGAAATGCTGAGCGCTAGAAATTCGACCAAAACATTTATTGAAATGG 660  
 QY 221 LeuAspGlnHisIleHisAlaTyrPheAlaValaGlyAlaProLeuLeuGlySerValGlu 240  
 Db 661 CTTTGATCAGCATATCCATGCTTATTTCGGTGTGAGCTCCCTCTTGTGTTCTGTGTAG 720  
 QY 241 AlaIleLysSerThrLeuSerGlyValaThrPheGlyLeuProValSerGluGlyThrAla 260  
 Db 721 GGAATCAAAATCACTCTCTGCTGTAAGCTTTGGCTTCCTGTTTCTAGGGAATCTGCT 780  
 QY 261 ArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCys 280  
 Db 781 CGGTGTTTTCCAATTCCTTTGGCTGCTGATGTGGCTTATGTCATTTCCAAAGAAATTC 840  
 QY 281 LysGlyAspAsnThrSerTrpThrHisPheSerGlyGlyAlaAlaLysLysAspLysArg 300  
 Db 841 AAGGGTGATACACATCCTGAGCGCATTTTTCGGGGGTGCGCAAGAAAGATTAAGCC 900  
 QY 301 ValTyrHisCysAspGluGluGlyTyrGlnSerLysTyrSerGlyTyrProThrAsnIle 320  
 Db 901 GATATACCACTGATGATAGAGAAATTCATCAAAATATTTCTGCGCGCAACAATATTT 960  
 QY 321 IleAsnIleGluIleProSerThrSerValThrGluThrAlaLeuValaAsnMetThrSer 340

Db 961 ATTAACATTGAATTCCTTCACCTAGCGTTACAGAAACAGCTTAGTCAACATGACCGAC 1020  
 QY 341 MetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThr 360  
 Db 1021 ATGGAATGTGGCTCTCCACCCCTTTTGTCTTCACGCCCGCTGAACATGACGATGGGACT 1080  
 QY 361 LeuPheLysAlaIleGluAspTyrAspProAspSerLysArgMetLeuHisGlyLeuLys 380  
 Db 1081 CTTTTCAAAGCAATAGAAAGCTATGACCCAGATAGCAAGAGATTTACACCAATTAAAG 1140  
 QY 381 LysLeuTyrHisAspAspProValPheAsnProLeuThrProTyrGluArgProIle 400  
 Db 1141 AAGTTGATCATATGATGACCTCTTTTATCTCTGACTCTTGGAGAGACCACTATA 1200  
 QY 401 LysAsnValPheCysIleTyrGlyAlaHisLysLysThrGluValaGlyTyrTyrPheAla 420  
 Db 1201 AAAAATGTATTTGGATATATGCTGCTCATCTAAAGACAGAGGTGGTGTATTACTTGGCC 1260  
 QY 421 ProSerGlyLysProTyrProAspAsnTrpIleIleThrAspIleIleTyrGluThrGlu 440  
 Db 1261 CCAGTGGCAACCTTATCTCGATATATGATTCATACGGATATCATTTAGCAAACTGAA 1320  
 QY 441 GlySerLeuValSerArgSerGlyThrValaLysGlyAsnAlaGlyProIleThrGly 460  
 Db 1321 GGTTCCTCGTGTCAAGGTCTGAACTGTGGTTGATGGAAAGCTGGACCTATACTGGG 1380  
 QY 461 AspGluThrValProTyrHisSerLeuSerTrpCysLysAsnTrpLeuGlyProLysVal 480  
 Db 1381 GATGAGCGGTACCTTATCATCTACTCTTTGGCAAGAAATTGGCTGGACCTAAAGTT 1440  
 QY 481 AsnIleThrMetAlaProGluProGluHisAspGlySerAspValaHisValaGluLeuAsn 500  
 Db 1441 AACATTAACAATGGCTCCCGACCCAGAACACATGAGAACGCGATCATGTGGAATTAAT 1500  
 QY 501 ValAspHisGluHisGlySerAspIleIleAlaAsnMetThrLysAlaProArgValLys 520  
 Db 1501 GTTGATCATGAGCATGGGTGACATCATCTACTTAACATGACAAAGACCAAGGTTAAG 1560  
 QY 521 TyrIleThrPheTyrGluAspSerGluSerIleProGlyLysArgThrAlaValaTrpGlu 540  
 Db 1561 TACATTAACCTTTTATGAAAGCTGTGAGACATTCGCGGGAAGAAACCGCAGTCTGGAG 1620  
 QY 541 LeuAspLysSerGlyTyr 546  
 Db 1621 CTGATTAAGAAGTGGAT 1638

RESULT 2  
 AY989885 1699 bp mRNA linear PLN 01-JUN-2005  
 LOCUS Arabidopsis thaliana phospholipid:sterol acyl transferase (PSAT)  
 DEFINITION  
 ACCESSION AY989885  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Arabidopsis thaliana (chale crese)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosoids; euroside II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1699)  
 Banaas, A., Carlsson, A.S., Huang, B., Lemman, M., Banaas, W., Lee, M.,  
 Noiriel, A., Benveniste, P., Schaller, H., Bouvier-Nave, P. and  
 Styenne, S.  
 Arabidopsis thaliana synthesizes sterol esters with an enzyme  
 related to the animal lecithin: cholesterol acyltransferase  
 Unpublished  
 2 (bases 1 to 1699)  
 Carlsson, A.S., Lemman, M., Lee, M. and Styenne, S.  
 Direct Submission  
 Submitted (30-MAR-2005) Crop science, Swedish University of  
 Agricultural Sciences, Sundsvagen 14, P.O. Box 44, Alnarp 23053,  
 Sweden  
 Location/Qualifiers

source 1. 1899  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
1. 21899  
/gene="PSAT"  
1. 21899  
/gene="PSAT"  
/function="catalyzes a transacylation of acyl groups from  
phospholipids to sterols"  
/product="phospholipid:sterol acyl transferase"  
/protein\_id="AA043920.1"  
/db\_xref="GI:66271903"  
/translation="MGANSKSVTASPTVIAFPFLICGGTAVEDTEPHDYSKLSGI  
IIIGRASPOLRAMSIIIDCPYPLDPMPLDWLDTTKLISAVNCFKCVLDPYDRLS  
HPECKRPDSGLSAILTELDPGYITGPISTVWKEMLKCVPEVEMALIVAPYDMRLS  
PKLEBRDLYFHKLKTFFETALKRGESIVFAHSMGNVFRILEMLKLEIAPKHYL  
KMLDHIHAYFAVGAFLGSVEAIKSTLSGVTFGLVPSGTAFLISNPFASLSLWMP  
SNCKGDNTPFTWTFSGAAKDKRVYHCDDEEYQSGYSGMPTNIIIEIPSTVETA  
LYNMTSMCEGLPTLSFTARELADGTLFAIDYDPSKRMHLQKLYHDPEVPL  
TPWERPIKNVFCIYGAHKTGVYFAFGSKPYDPMNTITDIIYETEGSLRSGLV  
VGNAGPTIGDETVPHSISCKNMLGPRVNTTMAPQPHDSDVHVELNDHEHSD  
IIANMTKAPRVKRTITFEDSESI PGKRTAVMELDKTNHKNIVRSPVLMRELWQMHND  
IOPGAKSFVTRAKRGLRDLADCCYMDYGRACCAQMOEYCEYRYSFGDVLGSGCLRNT  
SANMLQYI"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.4e-244 Length: 1899  
Score: 2944.00 Matches: 541  
Percent Similarity: 99.63% Conservative: 3  
Best Local Similarity: 99.08% Mismatches: 2  
Query Match: 99.29% Indels: 0  
DB: 15 Gaps: 0

US-09-651-651-5 (1-546) x AY989885 (1-1899)

QY 1 MetGlyAlaAsnSerIysSerValThrAlaSerPheThrValIleAlaValPhePheLeu 20  
DB 1 ATGGAGCGAATTCGAATCAGTAAGGCTTCTTCAACGCTCATCGCCGTTTTTTCTTG 60  
QY 21 ILeCysGlyValArgThrAlaValGluAspGluThrGluPheHisGlyAspTyrSerLys 40  
DB 61 ATTTGGCGTGGCGGAACGGCGGTGAGAGATGAGACCGAGTTTCACGGCGACTCGAAG 120  
QY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGluLeuArgAlaTTPSerIleLeu 60  
DB 121 GTATCGGATATAATCAATTCGGGATTTGCGTCAACGACGACTACGAGCGTGTGCATCTT 180  
QY 61 AspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValTTPLeuAspThrThr 80  
DB 181 GACTGTCATACCTCGTTGACCTTCAATCCGCTGCACTCGTATGCTAGACACCACT 240  
QY 81 LysLeuLeuSerAlaValAlaAsnCySTPPhelYsCysMetValLeuAspProTyrAsnGln 100  
DB 241 AACCTCTTTCTGCTGCACTGCTGTTTAAGTATGTGTGCTAGATCCTTAATCA 300  
QY 101 ThrAspHisProGluCysLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeu 120  
DB 301 ACAGACCATCCCAAGTGTAAGTCAACGGCTGACAGTGTCTTTCAAGCCATCACAGAAATTG 360  
QY 121 AspProGlyTyrIleThrGlyProLeuSerThrValTTPLeuGluTTPLeuLysTTPCys 140  
DB 361 GATTCACGTTACATTAACAGTCTCTTTTCTACTGTCTGGAAGAGTGGCTTAAGTGTGT 420  
QY 141 ValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAspTTPArgLeuSer 160  
DB 421 GTTGAGTTGTGTGAGAGCAAAATGCAATGTGCTGCTTCATACGATGGAAGATGTGCA 480  
QY 161 ProThrLysLeuGluGluArgAspLeuTyrPheHisLysLeuLysLeuThrPheGluThr 180

DB 481 CCAACCAAAATTGGAGAGCGTACCTTATCTTCAACAGCTCAAGTCACTTGAACACT 540  
QY 181 AlaLeuLysLeuArgGlyGlyProSerIleValIlePheAlaHisSerMetCysLysAsnVal 200  
DB 541 GCTTAAAACTCCGTGGCGGCCCTTCTATAGATTATGCCCATTCATGAGGTATATATGTC 600  
QY 201 PheArgTyrPheLeuGluTTPLeuArgLeuGluIleAlaProLysHisGlyTyrLeuLysTTP 220  
DB 601 TTCAGATACCTTCTGAAAGGCTGAGGCTAGAAATTGCACCAAAACATTAATTGAAAGTGG 660  
QY 221 LeuAspGlnHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuLysSerValGlu 240  
DB 661 CTGTGATCAGCATATCCATCTTATTCGTGTGAGGCTCTCTTCTGTGCTGTGAG 720  
QY 241 AlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGluGlyThrAla 260  
DB 721 GCAATCAAAATCACTCTCTGCTGTAACGTTTGGCTTCGTTTCTGAGGAACTGCT 780  
QY 261 ArgLeuLeuSerAsnSerPheAlaSerSerLeuTTPLeuMetProPheSerLysAsnCys 280  
DB 781 CGGTGTGTGTCAAATCTTTTCTGCTGCTATGTGCTTATGCCATTTTCAAGAAATTC 840  
QY 281 LysGlyAspAsnThrSerTTPThrHisPheSerGlyValAlaAlaLysLysAspLysArg 300  
DB 841 AAGGTGATTAACATTCCTGGAAGCATTTTCTGGGGGTGCTGCAAGAAAGATAGCCG 900  
QY 301 ValTyrHisCysAspGluGluGlyTyrGlnSerLysTyrSerGlyTTPProThrAsnIle 320  
DB 901 GTATACCACTGATGATCAAGAGAAATATCAATCAAAATATTCGCGCGCCGCAAAATATT 960  
QY 321 IleAsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsnMetThrSer 340  
DB 961 ATTAACATTGAATTCCTTCCATGCGTTACAGAAACGCTTACTACATGACACACAG 1020  
QY 341 MetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThr 360  
DB 1021 ATGGAAATGGCTCCCAACCTTTTGTCTTTCACAGCCCGTGAATACAGATGGAGACT 1080  
QY 361 LeuPheLysAlaIleGluAspTyrAspProAspSerLysArgMetLeuHisGluLeuLys 380  
DB 1081 CTTTCAAGCAATTAAGATCTATGACCCAGATAGCAAGAGATGTATACCCAGTTAAAG 1140  
QY 381 LysLeuTyrHisAspAspProValIlePheAsnProLeuThrProTTPGluArgProPheIle 400  
DB 1141 AAGTTGTATCATGATATCACTGTTTAAATCCTGTACCTCTGGAGAGACCACTATTA 1200  
QY 401 LysAsnValPheCysIleTyrGlyAlaHisLysLysThrGluValGlyTyrTyrPheAla 420  
DB 1201 AAAAATGATTTTGCATATATGTGCTCATTAAGACAGAGTGTGTTATTACTTGGC 1260  
QY 421 ProSerGlyLysProTyrProAspAsnTTPIleIleThrAspIleIleTyrGluThrGlu 440  
DB 1261 CCAAGTGGCAAACTTATCTGATTAATGATCATACCGAGATATCATTTATGAAACTGAA 1320  
QY 441 GlySerLeuValSerArgSerGlyThrPheValAlaAspGlyAsnAlaGlyProIleThrGly 460  
DB 1321 GGTTCCTCGTGTCAAGGTCTGAACTGTGTGAAGTGAAGGAACTGAACTTAATCTGGG 1380  
QY 461 AspGluThrValProTyrHisSerLeuSerTTPCysLysAsnTTPLeuGlyProLysVal 480  
DB 1381 GATGAACGGTACCTTATCACTCTCTGTTGCAAGATTTGGCTCGGACTTAAGTT 1440  
QY 481 AsnIleThrMetAlaProGlnProGluHisAspGlySerAspValHisValGluLeuAsn 500  
DB 1441 AACATTAACAATGCTCCCGCCAGCAACAGATGGAAGCACTAATGGAATCTAAT 1500  
QY 501 ValAspHisGluHisGlySerAspIleIleAlaAsnMetThrLysAlaProGlyValLys 520  
DB 1501 GTTGATCATGAGATGAGTCAACATCATATGATCAACAAAGACCAAGGCTTAAG 1560  
QY 521 TyrIleThrPheTyrGluAspSerGluSerIleProGlyLysArgThrAlaValTTPGlu 540  
DB 1561 TACATTAACCTTTTATGAAGCTGTGAGAGCATTCGGGGGAAGAAACCGCAGTCTGGAG 1620

QY 541 LeuAspLysSerGlyTyr 546  
 DB 1621 CTTGATMAAACAAATCAC 1638

RESULT 3  
 BD271623

LOCUS BD271623 3896 bp DNA linear PART 17-UNL-2003  
 DEFINITION Novel class enzyme in biosynthesis pathway of triacylglycerol  
 production and recombinant DNA molecule encoding the enzyme.  
 BD271623  
 ACCESSION BD271623.1 GI:33081391  
 VERSION BD271623  
 KEYWORDS JP 2002541783-A/8  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana (thale cress)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 3896)  
 Dahlqvist, A., Stahl, U., Lemman, M., Banas, A., Rönne, H. and Stymne, S.  
 Novel class enzyme in biosynthesis pathway of triacylglycerol  
 production and recombinant DNA molecule encoding the enzyme  
 Patent: JP 2002541783-A 8 10-DEC-2002;  
 BASF PLANT SCIENCE GMBH  
 OS Arabidopsis thaliana (thale cress)  
 PN JP 2002541783-A/8  
 PD 10-DEC-2002  
 PE 28-MAR-2000 JP 20060609586  
 PR 01-APR-1999 EP 99106656.4, 10-JUN-1999 EP 99111321.8  
 07-FEB-2000 US 60/180687  
 PI ANDERS DAHLQVIST, ULF STAHL, MARIT LEMMAN, ANTONI BANAS PI  
 HANS RÖNNE, STEEN STYMNE  
 PC C12N15/09, A01H5/00, C12N1/19, C12N5/10, C12N9/10, C12P7/64// PC  
 (C12P7/64, C12R1:645), (C12P7/64, C12R1:91), C12N15/00, C12N5/00 CC  
 Novel class enzyme in biosynthesis pathway  
 of triacylglycerol  
 CC production  
 CC and recombinant DNA molecule encoding the enzyme FH Key  
 FT Location/Qualifiers  
 FT source 1..3896  
 FT /organism='Arabidopsis thaliana (thale cress)'  
 FT location/Qualifiers  
 1..3896  
 /organism='Arabidopsis thaliana'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:3702'

Alignment Scores:  
 Pred. No.: 7e-191 Length: 3896  
 Score: 2332.00 Matches: 541  
 Percent Similarity: 41.584 Conservative: 0  
 Best Local Similarity: 41.584 Mismatches: 5  
 Query Match: 78.654 Indels: 759  
 DB: 6 Gaps: 11

US-09-651-651-5 (1-546) x BD271623 (1-3896)

QY 1 MetGlyAlaAsnSerLysSerValThrAlaSerPheThrValIleAlaValPhePheLeu 20  
 DB 1 ATGGGAGCGAATTCGAAATCAGTAACGGCTCTCTTACCGCTCATCGCGCTTTTCTTGG 60  
 QY 21 IIECGEGLYGlyARGThrAlaValGluAspGluThrGluPheHisGlyAspLys 40  
 DB 61 ATTTGGGTGGCGCGAATCGCGGTGGAGATGAGACGAGTTTCAACGCGCATCTCGAAG 120  
 QY 41 LeuSerGlyIleIleIlePheGlyPheAlaSerThrGluLeuArgAlaTrpSerIleLeu 60  
 DB 121 CTATCGGGATTAATCAATCTCGGGATTGCGTCGACGAGCTACGAGCGGTGATCTT 180  
 QY 61 AAspCyProTyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrThr 80

DB 181 GACTGTGCATACACTCCGTTGGACTTCAATCCGCTCGACCTCGTATGCGTAGACCACT 240  
 QY 81 Lys----- 81  
 DB 241 AAGGTCCGTGATCTTCAATTCCTTCGCTCTTAATTGTGCGGTGAGTCACTTGTGTATG 300  
 QY 81 ----- 81  
 DB 301 AATTCCAAAGCAAAATATAGCAATGAGCATGTCTGCTCTTAATTGATTTCGTTATTA 360  
 QY 81 ----- 81  
 DB 361 GTCAACAGTGAAGCTTGTGAATCTGAGTTTGAAGTCATATAAACAGCTGACCGGAG 420  
 QY 81 ----- 81  
 DB 421 TGTTCCTCATGCTTTGGTTCGCTAAATGTAGCGCAATGAATGTGTAATTAGTCTGGCC 480  
 QY 81 ----- 81  
 DB 481 TTTTATTCACTAGATCTGCAAGTTTTCAGAGTCTCATAGTATGTAAGAAATGTTA 540  
 QY 81 ----- 81  
 DB 541 GGTCAATTTTACTTGTGATGATGATCTTTGTTGTTGTTACTGATCGACGTGATGGA 600  
 QY 82 -----LeuLeuSerAlaValAsnCysTrpPheLysCysMetValLeuAspProTyr 98  
 DB 601 TGGTTTACAGCTCTTCTTCTGCTGTCACGTGCTGCTTGAAGTATAGGTGCTAGAACCTTA 660  
 QY 98 rAsnGlnThrAspHisProGluCysLeuSerArgProAspSerGlyLeuSerAlaIleThr 118  
 DB 661 TTAATMAACAGACCATCCCGAGTGTAGTCAAGCGGCTGACAGTGTCTTTCAGCATCAC 720  
 QY 118 rGluLeuAspProGlyTyrIleThr----- 126  
 DB 721 AGAATTGATCCAGGTACATTAAC-AGGTAGTTTCGATTTTCTTCTTTGAGTTTTC 779  
 QY 126 ----- 126  
 DB 780 TTCAATTGTGATATCATCTTGTGATGATATATATGCTTAAGTTCAATTTGGTCAATT 839  
 QY 127 -----GlyProLeuSerThrValTrpLysGlyLeuTrpLeuSerCysValGluPheGly 145  
 DB 840 TTCAGGTCTCTTCTTCACTGCTGGAAGAGGCGCTTAAGTGTGTGATTTGGTAT 899  
 QY 145 eGluAlaAsnAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuG 165  
 DB 900 AGAAGCAATGCAATGTCGCTGTTCCATATGAGATTGATGTCACCAACCAAAATTGGA 959  
 QY 165 uGluArgAspLeuTyrPheHisLysLeuLys----- 175  
 DB 960 AAGAGGTGACCTTTACTTTCACAAAGCTCAAGTTAGTCTTATCAGGCTAATGTCTTTTAT 1019  
 QY 176 -----LeuThr 177  
 DB 1020 CTTCTCTTTTATGTAAGATAAGCTAAGAGCTCTGTGCTCTCTTTTTCAGAGTTGAC 1079  
 QY 177 rPheGluThrAlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHisSerMetG 197  
 DB 1080 CTTTAAATCGCTTTTAAATCTCGTGGCGGCCCTTCTATAGTATTTGGCCATTCATGGG 1139  
 QY 197 yAsnAsnValPheArgTyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisIle 217  
 DB 1140 TAAATATGCTTCAATCTTCTGAAATGCTGAGGCTGAAGAAATTGCACCAAAACATTA 1199  
 QY 217 rLeuLysTrpLeuAspGlnHisIleHisAlaTyrPheAlaVal----- 231  
 DB 1200 TTTGAAGTGGCTTGAATCAGCAATATCATGCTTAATTTGCGTGT-TGGTACCGGCTAGTAT 1258  
 QY 231 ----- 231  
 DB 1259 CCTTAAGTTACATTTATTTTCTCTAATTTGGGGAGTTATGTGTGACTTACTGAT 1318



QY 232 -----GlyAlaProLeuLeuGlySerValGlu 241  
 DB 1319 TGAGTCGATACCTGATTTGTTGATTAGAGGCTCTTCTTGTTAGGCTCTGTTAGGCG 1378  
 QY 241 aileysSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu----- 257  
 DB 1379 AATCAATCTACTCTCTGCTGTAAAGTTGGCTCTGTTCTGAGGTGAAGCTCTGA 1438  
 QY 257 ----- 257  
 DB 1439 CTTCCTTTAGTTTAAAGTACTGATATCAACAGCTTTATTAACCTGATGATTTTCT 1498  
 QY 257 ----- 257  
 DB 1499 TTTGAAGTATTACTTTTGTAAATGAACTGCTGACGAGATATGGATCTGATGATCTT 1558  
 QY 257 ----- 257  
 DB 1559 GAAGTCTAGTATCAAGAACAATATTGGGTAGTATACCTGTCAAGCGCCTTAGCTAA 1618  
 QY 257 ----- 257  
 DB 1619 TACAACCAACACANGTACAGTGAATTTAGTTTACAGATTATTAATGAGATTTAAAGTT 1678  
 QY 257 ----- 257  
 DB 1679 GAGAGAAACTTGACTGAATCTTTTATTTTAAATAGGCTATGATTTGTTATTTGAAT 1738  
 QY 258 -----GlyThrAl 260  
 DB 1739 CATGTGACATATTGACATGCGCTTCTCATGTTTTTTTGTGGCAAGCCTTAGGGAACCTGC 1798  
 QY 260 aargLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLeuAsnCy 280  
 DB 1799 TCGGTGTGTTCACATTTCTTGTGCTGCTCATGTGTGCTTATGCTATTTTCAAGAAATG 1858  
 QY 280 slyGlyAspAsnThrSerTrpThrHisPheSerGlyAlaAlaLeuLysAspLysAr 300  
 DB 1859 CAAGGCGATATACCATTTCTGGAGCATTTTTTCTGGGGGTGCTCAAGAAAGATTAAGCG 1918  
 QY 300 gValTYHisCyAspGluGluGlyTYGlnSerLeuTYrSerGlyTrpProThrAsnI 320  
 DB 1919 CGTATACCACTGTGATGAAGGAAATATCATCAAAATATTTCTGGTGGCGCAAAATAT 1978  
 QY 320 eileAsnIleGluIleProSerThrSer----- 329  
 DB 1979 TATTAACTTGAATTCCTTCCACTAG-CGGTTAGACTCTGTATATGCAACTGTAACT 2037  
 QY 329 ----- 329  
 DB 2038 AACAAAGTTTCAACCAAGATGTTCACTCATATTTCCGTTCTTGATGTATCCATC 2097  
 QY 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCyGlyLeuProThrLeu 349  
 DB 2098 AGTTACAGAAACACCTCTAGTCAACATGACCAGCATGGAATGTGGCTTCCACACCTTT 2157  
 QY 349 uSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLysAlaIleGluAspTYrAs 369  
 DB 2158 GTCTTTTCAACGCCGTGTGAATGACAGATGGGACTCTTTTCAAGCAATATGAAGACTATGA 2217  
 QY 369 pProAspSerLysArgMetLeuHisGlnLeuLysLeu----- 382  
 DB 2218 CCCAGATAGCAAGAGATGTTACACCACTTAAAGAGTA-CGTACCTTTCTTTGTGATTA 2276  
 QY 382 ----- 382  
 DB 2277 GAAATATTGCTCATGATCATCACTGCTGGCTTCTGTAGCTGAATTTGTTTAA 2336  
 QY 382 ----- 382  
 DB 2337 ATCTGTATATCAATTGTTTCATATGCTTTGTCTTTCTTACTATTAAGAAACAAGTATATCA 2396

QY 382 ----- 382  
 DB 2397 GAAACCTTATTAATGATTAATACAGTCTCTCTTAATTAATGAAGTCTTTTCCGTTAC 2456  
 QY 382 ----- 382  
 DB 2457 AGTTATGAATGCAAAAGGGGATTTTACTGATTTGATTTCTCATTTCTAGTTTGT 2516  
 QY 382 ----- 382  
 DB 2517 TGACTAATAGCGTCAATTTGTTTCTAGCAAACTTTGTGAATTAATATAACATGCT 2576  
 QY 383 -----TyrHisAspAspProValPheAsnProLeuThrProTrpG 396  
 DB 2577 AACTATACCTTTTCAAGGTGTGATCATGATGACCTGTTTAAATCTTGACCTCTGGGA 2636  
 QY 396 uArgProProIleLysAsnValPheCysIleTYrGlyAlaHisLeuLysThrGlu----- 414  
 DB 2637 GAGACCACTATTAATAAATGTATTTTGCATATATGCTGCTCATTAAGACAGAGGTATG 2696  
 QY 414 ----- 414  
 DB 2697 ATGCATTTCOAATATCATATTAATGCTGATTTGTTATTAATTCCTCATTTGTTTGC 2756  
 QY 414 ----- 414  
 DB 2757 AATATCTTTTGAATTATGATTAATCTCTCCCTTGATCTTAATGCTTAAGCTTAA 2816  
 QY 415 -----ValGlyTYrTYrPheAlaProSerGly 424  
 DB 2817 GGTACTAAATATATGAAGCTGTCTGCATAGTGTGTTATTTACTTTGCCCAAGTGCGCA 2876  
 QY 424 sProTYrProAspAsnTrpIleIleThrAspIleIleTYrGluThrGluGlySerLeuVa 444  
 DB 2877 ACCTTATCCGATTAATTTGATCATCAGGATATCATTTATGAAGTGAAGTTCCTCGT 2936  
 QY 444 lSer----- 445  
 DB 2937 GTCAGAGTAATTTTCCGCAATGGCAGAGTAAGTAAACAGAAAGCTTCTGTATCAG 2996  
 QY 445 ----- 445  
 DB 2997 TCTAGTGCATGTTATCTCAGTTGCATAGCAAAATTTATTAACAATTAATTAAGTAC 3056  
 QY 445 ----- 445  
 DB 3057 TTTTATATCATCTCTTTTGAAGTTAGTATGATCAGTGGCTTAAAGTGGGAAGGTGT 3116  
 QY 445 ----- 445  
 DB 3117 TGCAAGAACATGACACTTGATATCAAAAGATTAAGTACAGCAAACTAACCCATTTCTG 3176  
 QY 445 ----- 445  
 DB 3177 AATTCATATTATTAAGAGTAGTCGCTTTTAAAAAATTTGTTTAAAGAACCGAAAA 3226  
 QY 446 -----ArgSerGlyThrValValAspGlyAsnAl 455  
 DB 3237 CTAGTTCATATCTTGATTTGCAATATCTGCAGGTCTGGAACCTGTGATGGGAAAGCG 3296  
 QY 455 aGlyProIleThrGlyAspGluThr----- 463  
 DB 3297 TGACCTTAATACGGGGATAGACGGTAACTCAGAAATTGGTTGAATAATCTTCTT 3356  
 QY 463 ----- 463  
 DB 3357 GCAAACTTACGAAGACTTAAGATTAATCTTCTTGTGAACACTGCTTGTATGTTCTTA 3416  
 QY 464 -----V 464  
 DB 3417 GTACACTGCAATATTAAGCTCTCCGCTACTTTTATTAATGAATGATCTTATAGG 3476  
 QY 464 aIProTYrHisSerLeuSerTrpCybLysAsnTrpLeuGlyProLysValAsnIleThrM 484

[illegible]

Score:	2332.00	Matches:	541
Percent Similarity:	41.58%	Conservative:	0
Best Local Similarity:	41.58%	Mismatch:	759
Query Match:	78.65%	Indel:	
DB:	6	Gaps:	11
US-09-651-651-5 (1-546) x BD271633 (1-3896)			
QY	1 MetGlyAlaenSerLysSerValThrAlaSerPheThrValIleAlaValPhePheLeu	20	
DB	1 ATGGAGCGAATTCGAATCAGTAACGGCTTCCTTCACCGTCATCGCGTTTTTCTTG	60	
QY	21 IleCysGlyGlyYarGThrAlaValGluAspGluThrGluPheHisGlyAspGlySerLys	40	
DB	61 ATTGCGGTGGCCGGAACGCGGTGGAGATAGACCGAGTTTCACGGGACATCTCGAAG	120	
QY	41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGluLeuArgAlaTrpSerIleLeu	60	
DB	121 CTATCGGGTATATATATTCGGGATTTTCGTCGACGCACTACGAGCGTGGTCGATCTT	180	
QY	61 AspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrThr	80	
DB	181 GACTGTCCATACACTCGGTGGACTTTCATCGCTCGACTGATGTAGCTAGACACT	240	
QY	81 Lys-----	81	
DB	241 AAGTCCGTGATCTTCATTCCTTCCTTCCTTATTCGTGGTCAGTCATCTGTTGATG	300	
QY	81 -----	81	
DB	301 AATTCACGCAATATATAGCAATGAAGCATGCTGCTCTCTTATGATTCGTTTCATTA	360	
QY	81 -----	81	
DB	361 GTCAACAGTAGACGTTCTGAATCTGAGTTTAGATCATATAAAGACTGACTCGCGAG	420	
QY	81 -----	81	
DB	421 TGTTTCCATGCGCTTTGGTGGCTGTAATGTAGCGCAATGAATGTGTAATGATCTGCGC	480	
QY	81 -----	81	
DB	481 TTTTATTCACATACTGCAGAGTTTTCAGAGTGCATATAGTAGTTAGAAATGTGA	540	
QY	81 -----	81	
DB	541 GGTCAATTTACTGTGTCATGTGATTCCTTTGGTGTGTTGCTTACTGATCGACGTGATGA	600	
QY	82 -----LeuLeuSerAlaValAsnCysTrpPheLysCysMetValLeuAspProTyr	98	
DB	601 TGGTTTACAGCTTCCTTTCGTGCTGCTACACGCTGGGTTTAAAGTATAGTGCTGATCTTA	660	
QY	98 rAsnGlnhTrAspHisrProGluCysLysSerArpProAspSerGlyLeuSerAlaIleThr	118	
DB	661 TAAATCAACAGACCAATCCGAGGTGAAGTCAGCGCTGACAGTGGTCTTTCAGCCATCAG	720	
QY	118 rGluLeuAspProGlyTyrIleThr-----	126	
DB	721 AGAATTTGATCCAGGTATCATTAAC-AGGTAGTTTGGATTTTCTTTGAGTTTTC	779	
QY	126 -----	126	
DB	780 TTCAAATTTGATTCATCTTGTGTGATATATATATGAGCTTAAGTTCAATTTGTCATAT	839	
QY	127 -----GlyProLeuSerThrValTrpLysGluTrpLeuLysTrpCysValGluPheGlyIle	145	
DB	840 TTCAGGTCTCTTCTTACTGCTGTGAACAGTGGCTTAAGTGTGTGTGAGTTTGAT	899	
QY	145 eGluhIAsnAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuGlu	165	
DB	900 AAGACCAATGCAATTTGTGGCTTCCATACGATTTGAGAAATTTGTCCAACCAATTTGGA	959	
QY	165 uGluArgAspLeuTyrPheHisLysLeuLys-----	175	

Db 960 AGAGCGTGAACCTTACTTTCACAGCTCAAGTTAGTCCTATCAGGCTAATGCTTTTAT 1019  
Qy 176 -----LeuH 177  
Db 1020 CTTCTCTTTTATGATAGATAGCTAAGAGCTCTGCTCGTCTTCTTTTTCAGGTTGAC 1079  
Qy 177 rPhegluThrAlaLeuLeuLeuArgGlyGlyProSerIleValPheAlaHisSerMetG 197  
Db 1080 CTTGGAACGCTTTTAAACCTCCGTCGGCCCTTCTATAGATTTGGCCATTCAATGGG 1139  
Qy 197 yAsnAsnValPheArgTyrPheLeuGluTyrPheLeuGluIleAlaProIlyshi sTy 217  
Db 1140 TAATTAATGCTCTCAGATCTTCTGGAATGCTGAGAGCTAGAAATTGCAACAAACATTA 1199  
Qy 217 rLeuLeuTyrPheLeuAspGlnHisIleHisAlaTyrPheAlaVal----- 231  
Db 1200 TTTGAAAGTGGCTGATCAGATACATACGCTTAATTTCGCTGT-TGGTACCGGCTTACTAT 1258  
Qy 231 ----- 231  
Db 1259 CCTTAAGTTACATTTTATTTTCTCTAATGGGGAGTTATGTTGTGACTTACGTGAT 1318  
Qy 232 -----GlyAlaProLeuLeuGlySerValGluAl 241  
Db 1319 TGAGCTGATACCTGATTTGTTGTGATTTAGGAGCTCCCTTCTTGCTTGAGGC 1378  
Qy 241 aAlIeLeuSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu----- 257  
Db 1379 AATCAATCTACTCTCTGCTGTAACGTTTGCGCTTCGTTCTGAGGTGACCTCTGA 1438  
Qy 257 ----- 257  
Db 1439 CTTCCTTAGTATTAGTAGTAGTATCAACAGGCTTATAACTGACTGATTTTCC 1498  
Qy 257 ----- 257  
Db 1499 TTTGAAGTATTACTTTTGTAAATTGAACGTCGTGACGATATAGTATCTGATCTT 1558  
Qy 257 ----- 257  
Db 1559 GAAGTCTAGTTATCAAGAACATATTTGGGTAATACCTGTCAAGGCTTAGCTAA 1618  
Qy 257 ----- 257  
Db 1619 TACAACCAACACATGTACATGATTAGTTTCAATTATTATGTAGACTTTAGTT 1678  
Qy 257 ----- 257  
Db 1679 GAGAGAACTTTGACTGAATCTTTTATTTTAAATAGGCTATGATTTGTTATTTGAAT 1738  
Qy 258 -----GlyThrAl 260  
Db 1739 CAGTGAATATGACATGCGCTTCTCATGTTTTTGTGGCAAGGCTTCAGGGAACGTC 1798  
Qy 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTyrPheLeuMetProPheSerIlyAsnCy 280  
Db 1799 TCGGTTGTTGCTCAATCTTTTGGCGTCAATGTGCTTATGCAATTTTCAAAAGATTG 1858  
Qy 280 sLyseGlyAspAsnThrSerTyrThrHisPheSerGlyGlyAlaAlaIlyAspIlySar 300  
Db 1859 CAAGGGATTAACATCTGACGCAATTTTCTGGGGGTGCTCCAAGAAAGATTAAGCG 1918  
Qy 300 gValTyrHisGlyAspArgGluGluTyrGlnSerLeuTyrSerGlyTyrProThrAsnI 320  
Db 1919 CGTATACCACTGATAGAGGAATCAATCAAAATATTCGTGCGCGCAAAATAT 1978  
Qy 320 eIleAsnIleGluIleProSerThrSer----- 329  
Db 1979 TATTAACTTGAATTCCTTCCACTAG-CGGTTAGACTCTGTATATGCAACTGTAACACT 2037  
Qy 329 ----- 329

Db 2038 AACAAAAGTTTACCACGAAGATGTCCTCATATTTCGTTCCTTGATGTATCATC 2097  
Qy 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCyseGlyLeuProThrLeuLe 349  
Db 2098 AGTTACAGAAACAGCTCTAGCTACATCATGACCGAGCATGGAAATGGCCCTTCCACCTTT 2157  
Qy 349 uSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheIlyValaIleGluAspTyrAs 369  
Db 2158 GTCTTTCACAGCCCGTGAACCTAGCAGATGGACTCTTTTCAAGCAATAGAAAGACTATGA 2217  
Qy 369 pProAspSerIlyAspArgMetLeuHisGlnLeuIlyAspIlyLeu----- 382  
Db 2218 CCGAGTACCAAGAGGATGTTACACAGTTAAAGAAAGTA-CGTACCTTTCTTTGTGATAA 2276  
Qy 382 ----- 382  
Db 2277 GAAATATGCTATGATCATCATCTGTGCTCTTGTAAGCTCAAAATTTGTTGTTAA 2336  
Qy 382 ----- 382  
Db 2337 ATCTATATCAATTGTTGATATGCTTGTCTTTCTTACTATPAGAAACAGTATATCA 2396  
Qy 382 ----- 382  
Db 2397 GAAACCTTATTTATGATTTACAGTCTCTCTTATATATGAAATGCTTTTTCGTTTAC 2456  
Qy 382 ----- 382  
Db 2457 AGTTATGAAATGCAAAAGGGGATATTTAGTGTATGATTCCTCATTCCTAGTTTGT 2516  
Qy 382 ----- 382  
Db 2517 TGACTAATAGCGTCAATTTTGTCTTCTAGCAAACTTTGTGATATATATAACATGCT 2576  
Qy 383 -----ThrHisAspAspProValPheAsnProLeuThrProThrPgl 396  
Db 2577 AACTATACCTTTTCAGGTTGTATCATATGACCTGTGTTTAACTCTGACTCTTGCGGA 2636  
Qy 396 uArgProProIleIlyAsnValPheCyseIleTyrGlyAlaHisLeuLeuThrGlu----- 414  
Db 2637 GAGACACATTAATAAAGATATTTTGCATATATGTGTGCTCATCTAAAGCAGAGGTATG 2696  
Qy 414 ----- 414  
Db 2697 ATGCATTCATATATCATATTAATGCTGTGACTTTGTTATATATCCCAATTTGTTGCT 2756  
Qy 414 ----- 414  
Db 2757 AATATCTTTTGAATTAATGATTTATCTCTCCCTTGCACTTATGCTATTAAGCGTTAA 2816  
Qy 415 -----ValGlyTyrTyrPheAlaProSerGlyLy 424  
Db 2817 GGTACTAAATGATGAAGCTGTCTGCATAGGTGTTATTTACTTTGCCCAAGTGCA 2876  
Qy 424 sProTyrProAspAsnTrpIleIleThrAspIleIleTyrGluThrGluGlySerLeuVal 444  
Db 2877 ACCTTATCCGATTAATTTGATCATCAAGATATCAATTTATGAAGCTGAAGTTCCCTGCT 2936  
Qy 444 lSer----- 445  
Db 2937 GTCAAGGTAAATTTTCCGCAATGCGCAAGTAAGTAACAGAAAGGCAAGCTTCTGTATCAG 2996  
Qy 445 ----- 445  
Db 2997 TCTAGTGCAATTTATCTCAGTTGCATTAAGCAATATTATTAACAATTAATTAAGTAC 3056  
Qy 445 ----- 445  
Db 3057 TTTTATATCATTCCTTTTAGCTTAGATGATCATGTGCTTAAAGTGAAGAGGTGT 3116  
Qy 445 ----- 445  
Db 3117 TGCATGAACATGACACTGTGATCAAAAGTAACTAGCAAAACAAACTAACCATTTCTG 3176

```

OY 445 ----- 445
Db 3177 AATTCAATATATTAGAGAGTAGTCGTCTTTTAAAAAATTTGTTTTAAGAAACCGAAAA 3236
OY 446 ----- 446
Db 3237 CTAGTTCATATCTTGATTGTGCAATATCTGCAGGCTCGAAGCTGTGGTTGATGGAGACCC 3296
OY 455 aglyProIleThrGlyAspGluThr----- 463
Db 3297 TGGACCTTAATACGGGGAGTAGACGGGTAAAGCTCAGAAATTGGTTTGAATTATCTTCTT 3356
OY 463 ----- 463
Db 3357 GCMAACTAGAGACTAAGATAATATCTTCTTGGAACACTGCTGATGTTCTTA 3416
OY 464 ----- 464
Db 3417 GTACACTGCATATTTGACTCTCGCTACTTTTATTGATTATGAATTGATCTTTATAG 3476
OY 464 aIProGlyHisSerLeuSerTyrCysIleAsnIlePheGlyProIleValAsnIleThrM 484
Db 3477 TACCCATCATTCACCTCTTGTTGTCAGAAATTGGCTCGAAGCTTAAGTTAACAATACAA 3536
OY 484 eGAla-Pro----- 486
Db 3537 TGGCTCCAGAGTACTCTTTTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3596
OY 486 ----- 486
Db 3597 TTTTCTGTTATGTGTGATTACTCTCAATTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3656
OY 487 ----- 487
Db 3657 TACTCTCAAGAACTTGATTAATCTAAACGAGATTCATGGAAGAAATTAACAAACAG 3716
OY 488 ProGluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySer 507
Db 3717 CCAGAACAGATGGAAGCGACGTACATGTGGAACCTAAATGTGATATGACATGGGCTCA 3776
OY 508 AspIleIleAlaAsnMetThrIleValAspArgValIleThrPheTyrGluAsp 527
Db 3777 GACATATACGTACACATGACAAAGACACCAAGGTTTACGTACATTAACCTTTATGAAGAC 3836
OY 528 SerGluSerIleProGlyIleValArgThrAlaValIlePheGluLeuAspIleSerGlyTyr 546
Db 3837 TCTGAGAGCATTCGGGGAAGAGAACGCGAGCTGAGAGCTGATTAAGTGGGAT 3893

RESULT 5
LOCUS AX037587 3896 bp DNA linear PART 16-NOV-2000
DEFINITION Sequence 11 from Patent WO060095.
ACCESSION AX037587
VERSION AX037587.1 GI:11227006
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
Banae,A., Stahl,U., Symme,S., Lemman,M., Ronne,H. and Dahlqvist,A.
A new class of enzymes in the biosynthetic pathway for the
production of triacylglycerol and recombinant dna molecules
encoding these enzymes
Patent: WO 0060095-A 11 12-OCT-2000;
BASF PLANT SCIENCE GMBH (DE) ; BANA ANTONI (PL) ; STAHL ULF (SE) ;
STYME STEN (SE) ; LEMMAN MARIT (SE) ; RONNE HANS (SE) ; DAHLQVIST
ANDERS (SE)
Location/Qualifiers
1..3896
/organism="Arabidopsis thaliana"

```

```

/mot_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 7e-191 Length: 3896
Score: 2332.00 Matches: 541
Percent Similarity: 41.58% Conservative: 0
Best Local Similarity: 41.58% Mismatches: 5
Query Match: 78.65% Indels: 759
DB: 6 Gaps: 11
US-09-651-651-5 (1-546) x AX037587 (1-3896)
OY 1 MetGlyIleAsnSerIleSerValThrIleSerPheThrValIleAlaValPhePheLeu 20
Db 1 ATGGAGGCGAATTCGAAATCAGTAACGGCTTCCTTACCGCATCGCCGTTTCTTG 60
OY 21 IleCysGlyGlyArgThrAlaValGluAspGluThrGluPheHisGlyAspTyrSerIle 40
Db 61 ATTTCGGTGGCCGGAACCTCGCGTGGAGATGAGACCGAGTTTCAAGGAGACTCGAAG 120
OY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGluLeuArgAlaTyrSerIleLeu 60
Db 121 CTATCGGGTATATCATTCCTCGGATTTGCGTCAAGCGAGCTTACGAGCGGTGCGATCCTT 180
OY 61 AspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValIlePheAspThrThr 80
Db 181 GACTGTCATACACACCGGTGAGCTTCAATCCGCTGACCTCGATGAGCTTACACCACT 240
OY 81 Iys----- 81
Db 241 AAGGTCGGGATCTTCATTTCTTCTGCTCTTATTCGTGCGTGCAGTCACTGTTGATG 300
OY 81 ----- 81
Db 301 AATTCGAAGCAATATAGCAATGAAGCATGTCGTCTCTTATGATTGTTGATTCATTA 360
OY 81 ----- 81
Db 361 GTCAACAGTACGCTTCTGATCTGAGTTTGAAGTCAATATAAACAAGCTGACTCGGCGAG 420
OY 81 ----- 81
Db 421 TGTTCCTCATGCGCTTTTGTTGCTGCTAATGTAGCGCAATGAATGTATTAGTGTGGCC 480
OY 81 ----- 81
Db 481 TTTTATTCACATGATCTGCAGAGTTTTCAGAGTGTCAATAGTAGTTAGAAAAATGTTA 540
OY 81 ----- 81
Db 541 GGTCAATTTTACTTGTCATTTGATTTCTTTGTTGTTGCTTACTGATCGACGTATGGA 600
OY 82 -----LeuLeuSerAlaValAsnCysIlePheIleCysMetValLeuAspProTyr 98
Db 601 TGGTTTACAGCTTCTTTCTGCTGCTCACTGCTGCTTAAAGTATGATGCTGATGCTTAA 660
OY 98 rAsnGlnThrAspHisProGluCysIleSerArgProAspSerGlyLeuSerAlaIleThr 118
Db 661 TATATCAACAGACCAATCCGAGGTGTAAGTCAAGGCTGACAGTGTGTTTCAGCATCAC 720
OY 118 rGluLeuAspProGlyTyrIleThr----- 126
Db 721 AGAATTGGATCCAGGTATCATAC-AGGTAGTTTGGATTTTCTTTCTTGACGTTTTC 779
OY 126 ----- 126
Db 780 TTCATATTGATATCATCTTGTGTGATATATATGAGCTTAAGTTCAATTAATTGTCATTT 839
OY 127 ---GlyProLeuSerThrValIlePheGlyIlePheLeuIlePheValGluPheGlyIle 145
Db 840 TTCAGTCTCTTCTTCTACTGTCTGGAAGAGAGGCTTAAGTGTGTGTGATGTTGGTAT 899

```

145 egluaiaaenalailevalaValProtyrAspTyrArgLeuSerProThrylsleuGl 165  
 900 AGAAGCAAAATGCAATTGTGGCTGTCATACATGAGATTGTACCAACCAAAATTGGA 959  
 165 uGluAAspLeuTyrPheHiblyleuLys----- 175  
 960 AGAGCTGACCTTACTTTCACAAAGCTCAAGTTAGTCCTTATCAGGCTAATGCTTTTAT 1019  
 176 -----LeuTh 177  
 1020 CTTCTCTTTTATATAGATAACTAAGACTCGTGGCTCTTCTTTTGCAGGTGAC 1079  
 177 rPheGluThrAlaLeuLysLeuArglyProSerIleValPheAlaHisSerMetGl 197  
 1080 CTTTGAACCTGCTTAAACTCCGTGGCGGCCCTTCTATATGATTTGGCCATTCATGGG 1139  
 197 yAsnAsnValPheArgTyrPheLeuGluTyrPheLysGluIleAlaProLysHisTyr 217  
 1140 TAATAATGCTTCAGATTAATCTTCTGGAATGCTGAGAGTGAAGAAATTCACAAACATTA 1199  
 217 rleuLysTyrPheLysPglHisIleHisAlaTyrPheAlaVal----- 231  
 1200 TTTGAAGTGGCTGATCAGATACATACCTTAATTGCTGT-TGGTACCGGCTTACTAT 1258  
 231 ----- 231  
 1259 CCTTAAGTACATTTTATTTTCTCTAATTGGGGAGTTATGTTGACTAGAT 1318  
 232 -----GlyAlaProLeuLeuGlySerValGluAl 241  
 1319 TGAGCTGACATCCGATTTGTTGTGATTTAGAGACTCTTCTGTTCTGTGAGGC 1378  
 241 aileLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu----- 257  
 1379 AATCAATCTACTCTCTGTGTGAAGCTTTGGCTTCTCTTTCTGAGGTGACCTCTGA 1438  
 257 ----- 257  
 1439 CTTCTCTTAGTTTAAAGTAGTGTGATATCAACAGGCTTATTAACCTAGATTTTCT 1498  
 257 ----- 257  
 1499 TTTGAAGTATTACTTTTGTTAATTGAACGCTGTACGCGATATGATCTGATCTT 1558  
 257 ----- 257  
 1559 GAAAGTCTAGTATCAAAAGACATATTGTGGGTATGATACCTGTCAAGGCTTAGCTAA 1618  
 257 ----- 257  
 1619 TACACCAAAACCATGTACATGATTTAGTTTCACATTATTATGTTAGACTTTAAGTT 1678  
 257 ----- 257  
 1679 GAGAGAAACTTGGACTGAATCTTTTATTTAATAGCTATGATTTGTTATTTGAAT 1738  
 258 -----GlyThrAl 260  
 1739 CATGTGACATATTTGACATGGCGCTTCATGTTTTTTTGGCAAGGCTTCAGGGAATGCG 1798  
 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTyrPheLysMetProPheSerLysAsnCy 280  
 1799 TCGGTTGTGTCCAAATCTTTTGGCTGTGCTATGTGCTTATGTCATTTTCAAGAAATTTG 1858  
 280 blyArgLysAspAsnThrSerTyrPheHisPheSerGlyValAlaAlaLysLysAspLysArg 300  
 1859 CAAGGCGATTAACCATTTCTGGAGCATTTTCTGGGGGTGCTGCAAAAGAAAGTAAAGCG 1918  
 300 gValTyrHisCyAspPglLysGluIleTyrGlnSerLysTyrSerGlyTyrProThrAsnIle 320  
 1919 CGTATACCACTGTATGAGAGGAATATCAATCAAAATATTCTGGCTGGCGCAAAATAT 1978

320 eileAsnIleGluIleProSerThrSer----- 329  
 1979 TATTAACATTTGAAATTCCTTCCACTAG-CGGTTAGACTCTGTATATGCACCTGTAACACT 2037  
 329 ----- 329  
 2038 AACCAAAAGTTTACCAAGAAATTTGACTCTCATATTGCTTCTTGAATGTATCCATC 2097  
 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGlyCyArglyLeuProThryleu 349  
 2098 AGTTACAGAAACAGCTCTAGTCACATGACCCAGATGGAATGTGGCTTCCACCCCTTTT 2157  
 349 uSerPheThrAlaArgGluLeuValAspGlyThrLeuPheLysAlaIleGluAspTyrAs 369  
 2158 GTCCTTACAGCCCGTGAACCTGACAGATGGACTCTTTTCAAGCAATAGAAAGACTATGA 2217  
 369 pProAspSerLysArgMetLeuHisGlyIleLysLysLeu----- 382  
 2218 CCCAGATTACAAAGAGATGTTACACCAGTTAAAGAAATG-CGTACCTTTCTTTGTGATGA 2276  
 382 ----- 382  
 2277 GAAATATGCTCATGATCATCATCTGTGCTTCTTGTACGTCAAAATTTGTTTAA 2336  
 382 ----- 382  
 2337 ATCTATATCAATTTGTCATATGCTTGTCTTCTTACTATAGAAACAGATATATCA 2396  
 382 ----- 382  
 2397 GAAACCTTATTTATGATATACATGTTCTCTCTTATATATGAAGTCTTTTTCGTTAC 2456  
 382 ----- 382  
 2457 AGTTATGAATGCMAAAGGGGATTTTAGTATGATTCCTCATCTCTAGTTTGT 2516  
 382 ----- 382  
 2517 TGACTAATAGCGTCAATTTTGTCTTCTAGCAAACTTTGTGAATATATATATACATGCT 2576  
 383 -----TyrHisAspAspProValPheAsnProLeuThrProThrPgl 396  
 2577 AACCTATCTTTTCAAGTGTGATATCATATGACCTGTTTTTAACTCTGACTCTTGGGA 2636  
 396 uArgProPheLysAsnValPheCysIleTyrGlyAlaHisLeuLysThrGlu----- 414  
 2637 GAGACCACTATTAATAAATGATTTTGCATATATGCTCATCTTAAGACAGAGGTATG 2696  
 414 ----- 414  
 2697 ATGCATTCATATATCATATATATGCGTGAATTTGTTATATATCCCATTTGGTTTGC 2756  
 414 ----- 414  
 2757 AATATCTTTTGAATATGATTTATCTTCTCCCTGCATCTTATGCTATTAAGCTTAA 2816  
 415 -----ValGlyTyrTyrPheAlaProSerGlyLys 424  
 2817 GGTACTAAATGATAGAGCTGTCTGTCAATGAGTTGTTATTAATCTTGGCCCAAGTGCA 2876  
 424 sProGlyTyrProAspAsnTyrIleIleThrAspIleIleTyrGluThrGlyLysLeuVal 444  
 2877 ACCTTATCCGATTAATTTGATATCATACGATATATCATTTATGAAGCTGAAGGTCCCTCGT 2936  
 444 lSer----- 445  
 2937 GTCAAGATTAATTTCCGCAATGCGAAGTAAGTAAGAAAGCAAGCAAGTCTTCTGTATCAG 2996  
 445 ----- 445  
 2997 TCTAGTGCATGTATCTCAGTTGACATAGCAATTTATTAACACTAAATTTAAGTAC 3056  
 445 ----- 445

Db 3057 TTTTTCATTCCTTTTGAAGCTAGTATGATGATGCTTAAAGTGAAGAGCTGT 3116  
QY 445 ----- 445  
Db 3117 TGCATGAACATGACACTGTGTATCAAGATTAAGCAAAACAACTAACCCATTCTTG 3176  
QY 445 ----- 445  
Db 3177 AATTTCATTTATTAGAGAGTACTGCTGCTTTTAAAAAATTTGTTTAAAGAAACCGAAAA 3236  
QY 446 ----- ArgSerGIYThrValValAspGIYAsnAl 455  
Db 3237 CTAGTTTCATATCTTGATTGTGCAATATCTGCAAGTCTGGAAGCTGTGATTGATGGGAACGC 3296  
QY 455 agIYProIIeThrGIYAspGIYThr----- 463  
Db 3297 TGGACCTTAATACGGGGATGAGACGGTAAGCTCAGAAATTGGTTTGAATTATCTTCTT 3356  
QY 463 ----- 463  
Db 3357 GCAAACTAGTAAGACTAAGATAATATTCTTCTGAAACACTGCTGTATGTTCTCTA 3416  
QY 464 ----- V 464  
Db 3417 GTACACTGCAATATTGACTCTCCGCTACTTTATTGATTATGAATTTGATCTTTATAG 3476  
QY 464 aIProIYrHisSerIeuSerIYrCylValAsnIYrPleuGIYProIYsValAsnIYrThM 484  
Db 3477 TACCCATATCATCTACTCTTGGTGCAAGAAATTGGCTCGAAGCTTAAAGTTAATACAA 3536  
QY 484 eEAIs-Pro----- 486  
Db 3537 TGGCTCCCAAGGTACTCTTTTATTTAGTTCTCCTCACTTATATAGTCAAACTTTAAGCTAC 3596  
QY 486 ----- 486  
Db 3597 TTTTTCGTTATGTTGATTACCTCAATTGTTCTTTCTAAATCATATATCTCG 3656  
QY 487 ----- Gln 487  
Db 3657 TACTCTCAAGAACTGTATTATTAATCTAAACGAGATTTCTACTGGGAAAAATAAACAAACG 3716  
QY 488 ProGIuHIsAspGIYSerAspValHisValGIuLeuAsnValAspHisGIuHIsGIYSer 507  
Db 3717 CCGAAGACAGATGAGACGACGTACATGTGAACTAAATGTGATCATGAGCATGGGTCA 3776  
QY 508 AspIleIleAlaAsnMetThrIYsAlaProArgValIYrIleThrPheYrGIuAsp 527  
Db 3777 GACATCATAGCTAACATGACAAAGACACCAAGGTTTAAGTACATACTTTTATGAAGAC 3836  
QY 528 SerGIuSerIleProGIYLYsArGIuThAlaValITrGIuLeuAspIYsSerGIYr 546  
Db 3837 TCTGAAGAGCATTCGGGGAGAGAAACCGCAGTCTGGAGCTGTATAAAGTGGGTAT 3893

RESULT 5  
AX037606 AX037606 3896 bp DNA linear PART 16-NOV-2000  
DEFINITION Sequence 30 from Patent WO060095.  
ACCESSION AX037606  
VERSION AX037606.1 GI:11227020  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE  
AUTHORS Baner, A., Stahl, U., Syme, S., Lemman, M., Ronne, H. and Dahlqvist, A.  
TITLE A new class of enzymes in the biosynthetic pathway for the  
production of triacylglycerol and recombinant dna molecules  
encoding these enzymes  
JOURNAL Patent: WO 0060095-A 30 12-OCT-2000;

FEATURES  
source  
Location/Qualifiers  
1..3896  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Alignment Scores:  
Pred. No.: 7e-191 Length: 3896  
Score: 2332.00 Matches: 541  
Percent Similarity: 41.58% Conservative: 0  
Best Local Similarity: 41.58% Mismatches: 5  
Query Match: 78.65% Indels: 759  
DB: 6 Gaps: 11

US-09-651-651-5 (1-546) x AX037606 (1-3896)

QY 1 MetGIYAlaAsnSerIYsSerValThrAlaSerPheThrValIleAlaValPhePheIeu 20  
Db 1 ATGGAGCGAATTGAAATTCAGTAACGGCTTCTTCCATCGATCGCCGTTTTCCTTG 60  
QY 21 IleCysGIYGIYArgThrAlaValGIuAspGIuThrGIuPheHisGIYAspIYsSerIYs 40  
Db 61 ATTTCGGTGGCCGCAACTCGCGGTGAGATGAGACGCAAGTTTACAGGGGACTACGAAG 120  
QY 41 IeuSerGIYIleIleIleProGIYLYsAlaSerThrGIuLeuAGAlaIYrSerIleIeu 60  
Db 121 CTATCGGGGTATATCATTCGCGGATTTTCGTGACGACGATACGAGCGGTGCTTCCTT 180  
QY 61 AspCysProIYrThrProIeuAspPheAsnProIeuAspIYrValITrPleuAspThr 80  
Db 181 GACTGCCATACACTCCGTTGGACTTCAATCCGCTCGAAGCTGTATGGCTAGACCACT 240  
QY 81 IYs----- 81  
Db 241 AAGGTCGGATCTTCATTTCTGCTCCTTATTCTGCGGTGAGTCACTTGTGATG 300  
QY 81 ----- 81  
Db 301 AATTCAGCGAATATAGCAATGAAAGATGTCTGCTCTTATGATTGCTTCATTA 360  
QY 81 ----- 81  
Db 361 GTCAACAGTGACGCTTCTGATCTGAGCTTGAAGTATGAATATAAACAGCTGACTCGCGAG 420  
QY 81 ----- 81  
Db 421 TGTTCGCCATGCGCTTTGGTTCGTAATGTAGCCCAATGAATGTATAGTCTGCGC 480  
QY 81 ----- 81  
Db 481 TTTTATTCACATAGATCTGCAAGTTTTCAGAGTCTCAATAGTAGTAAATGTTA 540  
QY 81 ----- 81  
Db 541 GGTCAATTTTACTTGGCATTTGATGATCTTTTGGTTGCTTACTGATCGACGTATGGA 600  
QY 82 -----LeuIeuSerAlaValAsnCIYrPheIYsCysMetValIeuAspProIY 98  
Db 601 TGGTTTACAGCTTCTTTCTGCTGCTCACTGGTTTAAAGTGTATGGGTCTAGATCTT 660  
QY 98 rAsnGIuThrAspHisProGIYLYsAlaSerArgProAspSerGIYLeuSerAlaIleTh 118  
Db 661 TATCAAAAGACCAATCCGAGGTGAAGTACAGGCGCTGACAGTGTCTTTTACGACATCAC 720  
QY 118 rGIuIeuAspProGIYrIleThr----- 126  
Db 721 AGAATGGATCCAGGTATACATAAC-AGGTAGTTTGGATTTTCTTTTGGAGTTTC 779  
QY 126 ----- 126

Db 780 TTCAATTTGATATCATCTTGTGTGATATATATGCTTAAGTTCAATTAATTTGGTCAATT 839  
 Qy 127 ----GlyProLeuSer-ThrValTrpLeuGluTrpLeuLysTrpCysValGluPheGlyIle 145  
 Db 840 TTCAGGATCCCTCTTCTACTGCTCTGGAAAGAGGCTTAAGTGCTGTTGAGTTGGTAT 899  
 Qy 145 egluaIaAenAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuGlu 165  
 Db 900 AGAAGCAAAATGCATTGTGCTGCTTCATACGATTGTGAGATTGTCCACCAACCAAAATTGA 959  
 Qy 165 uGluArgAspLeuTyrPheHisLysLeuLys----- 175  
 Db 960 AGAGCTGACCTTTACTTTACACAGCTCAAGTTAGCTTATATCAGGCTAATGCTTTTAT 1019  
 Qy 176 -----LeuThr 177  
 Db 1020 CTTCTCTTTTATATGATAGATTAAGCTTAAGACCTCGTCTCTTTCTTTTGACAGTTGAC 1079  
 Qy 177 rPheGluThrAlaLeuLysLeuArgGlyProSerIleValPheAlaHisSerMetGlu 197  
 Db 1080 CTTTGAACCTGCTTTAAACTCCGTCGCGCCCTTCATATAGTATTTGCCATTCAAATGGG 1139  
 Qy 197 yAsnAsnValPheArgTyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisTyr 217  
 Db 1140 TTAATAATGCTTCAGATACTTCTCGAATGCTGAGGCTAAGAAATTGCACCAAAACATTA 1199  
 Qy 217 rLeuLysTrpLeuAspGluHisIleHisAlaTyrPheAlaVal----- 231  
 Db 1200 TTTGAAGTGGCTGATCAGCATATCCATGCTTATTTGCTGT- TGGTACCGGCTACTAT 1258  
 Qy 231 ----- 231  
 Db 1259 CCTTAAGTTACATTTTATTTTCTCTAATTGGGGAGTTATGTTGACTTACTGAT 1318  
 Qy 232 -----GlyAlaProLeuLeuGlySerValGluAla 241  
 Db 1319 TGAGCTGATACCTGATTTGTGTGATTTAGAGAGCTCCCTCTTGCTGTCTGTGAGGC 1378  
 Qy 241 aIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu----- 257  
 Db 1379 AATCAATCTACTCTCTGCTGTGATGAGCTTTGGCTTCTGTTCTTGAGGTGACCTCTGA 1438  
 Qy 257 ----- 257  
 Db 1439 CTTCTCTTAGTTTAAAGTAGTGTGATATCAACAGGCTTATTAACCTACGTGATTTTCT 1498  
 Qy 257 ----- 257  
 Db 1499 TTTGAAGTATTACTTTTGTAAATTGAACGCTGTACGCGATATGATATCTGATCTT 1558  
 Qy 257 ----- 257  
 Db 1559 GAAAGTCTAGTTATCAAGAACATATTTGGGTATGATACCTGTACGGGCTTAGCTAA 1618  
 Qy 257 ----- 257  
 Db 1619 TACAACCAACACATGTACATGATTTAGTTTACATTAATTATGTGTAGACTTAAAGTT 1678  
 Qy 257 ----- 257  
 Db 1679 GAGAGAAACCTTGACTGAATCTTTTATTTTAAATAGCTATGATTTGTTTAAAT 1738  
 Qy 258 -----GlyThrAla 260  
 Db 1739 CATGTGACATATTGACATGCGCTTCTCATGTTTTTTTGTGGCAAGGCTTCAGGAACTGC 1798  
 Qy 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnGly 280  
 Db 1799 TCGGTTGTGTGCCAATCTTTTGTGCTGTATGCTATGCTATGCTATTTCAAGAAATTG 1858  
 Qy 280 sLysGlyAspAsnThrSerTrpThrHisPheSerGlyValAlaAlaLysAspLysArg 300

Db 1859 CAAGGTGATTAACACATTTCTGACCGCATTTTCTGGGGGTGTGTCGAAGAAAGATMAAGC 1918  
 Qy 300 gValIyTrHisCysAspGluGluGluTrpGlnSerLysTyrSerGlyTrpProThrAsnIle 320  
 Db 1919 CGTATACACACTGTGATGAAGAGGAATATCAATCAAAATATCTGCTGGCCGACAAATAT 1978  
 Qy 320 eIleAsnIleGluIleProSerThrSer----- 329  
 Db 1979 TATTAACATTTGAATTCCTTCCACTAG- CGGTTAACCTCTGTATATGCACCTGTAACT 2037  
 Qy 329 ----- 329  
 Db 2038 AACCAAAAGTTTACCAAGAAATGTCACTCTCATTTTGTGCTCTTGAATGTATCCATC 2097  
 Qy 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeuLe 349  
 Db 2098 AGTTACAGAAACAGCTCTGATCAACATGACCCAGCATGAGATGTGCTCTTCCACCTTTT 2157  
 Qy 349 uSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLysAlaIleGluAspTyrAs 369  
 Db 2158 GTCTTTCACAGCCCGTGAACATGACAGATGGAGCTCTTTCAAGCAATAGAACTATATGA 2217  
 Qy 369 pProAspSerLysArgMetLeuHisGluLeuLysLysLeu----- 382  
 Db 2218 CCCAGATTGCAAGAGATGTTACACCAGTTAAAGAAAGTA- CGTACCTTTCTTTGTGATAA 2276  
 Qy 382 ----- 382  
 Db 2277 GAAATATGCTCATGATCATCATCTGTGCTCTGTGACGTCAAAATGTTTGTATTA 2336  
 Qy 382 ----- 382  
 Db 2337 ATCTATATCAATTTGTTATGATATGCTTGTCTTTCTTACTATAGAAACAAGATATATCA 2396  
 Qy 382 ----- 382  
 Db 2397 GAAACTTATTTATGATTTACAGTCTCTCTTATATATGAAAGTCTTTTTCGTTTAC 2456  
 Qy 382 ----- 382  
 Db 2457 AGTTATGAATGCMAAGGGGATTTTATGATGATGATTTCTCATTTCTAGTTTGT 2516  
 Qy 382 ----- 382  
 Db 2517 TGACTAATAGCGTCAATTTGTTTTCTAGCAAACTTTGTGAATATATATATACATGCT 2576  
 Qy 383 -----TyrHisAspAspProValPheAsnProLeuThrProTrpGlu 396  
 Db 2577 AACTATACCTTTTCAGGTTGTATCATATGATGACCTGTGTTTAACTCTGACCTCTGGGA 2636  
 Qy 396 uArgProProIleLysAsnValPheCysIleTyrGlyAlaHisLysLeuThrGlu----- 414  
 Db 2637 GAGACCACCTATPAAAAATGATTTTGTGATATGTGCTCATCTPAAAGACAGAGGTATG 2696  
 Qy 414 ----- 414  
 Db 2697 ATGCATCTCAATATCATATATGCTGTGACTTTGTATTAATATCCCAATTGGTTGCT 2756  
 Qy 414 ----- 414  
 Db 2757 AATATCTTTTGAATTAATATGATTTATCTTCTCCCTTGACCTTAATGCTATTAAGCTTAA 2816  
 Qy 415 -----ValGlyTyrTyrPheAlaProSerGlyLys 424  
 Db 2817 GGTACTAAATGTATGAAGCTGTCTGATATAGTTGTGTTATCTTTGCCCCAAGTGCAG 2876  
 Qy 424 sProTyrProAspAsnTrpIleIleThrAspIleIleTyrGluThrGluGlySerLeuVal 444  
 Db 2877 ACCTTATCCGATTAATTTGATATCAAGATATCATTTATGAAGATCGAAGTTCCCTCGT 2936  
 Qy 444 lSer----- 445  
 Db 2937 GTCAAGTAAATTTTCCGCAATGCGCAAGATGAAGAAAGCAAGAGTCTTCTGTATCAG 2996



[illegible]

REFERENCE	1 (bases 1 to 119914)
AUTHORS	Federapfel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,I., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetakala,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoro,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaka,V.S., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Unpublished
TITLE	2 (bases 1 to 119914)
REFERENCE	Federapfel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R., Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y., Li,O., Osborne,B.I., Shinn,P., Sun,H., Tortum,M., Vyotskaka,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Direct Submission
TITLE	Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	3 (bases 1 to 119914)
AUTHORS	Federapfel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,I., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetakala,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoro,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaka,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Direct Submission
TITLE	Submitted (30-DEC-1998) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE	4 (bases 1 to 119914)
AUTHORS	Federapfel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Araujo,R., Huizar,I., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetakala,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoro,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaka,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Direct Submission
TITLE	Submitted (30-JAN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Dec 30, 1998 this sequence version replaced g1:2734094. Bases 1-9262 of clone F21M1 overlap with bases 68998-78259 of 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone F21M1 overlap with bases 1-389 of 'TAMU' BAC clone F21B7 (AC002560).
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu
COMMENT	Genes with similarity to proteins in the databases are described as 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <a href="http://combio.ornl.gov/section/index.html">http://combio.ornl.gov/section/index.html</a> ), GENSCAN (Chris Burge, <a href="http://genomic.stanford.edu/~chris/GENSCANW.html">http://genomic.stanford.edu/~chris/GENSCANW.html</a> ), Fexa (V.Solovayev & A.Salamov, Sanger Centre, <a href="http://genomic.sanger.ac.uk/">http://genomic.sanger.ac.uk/</a> ), and NePlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <a href="http://www.cbs.dtu.dk/NePlantGene.html">http://www.cbs.dtu.dk/NePlantGene.html</a> ).
FEATURES	Location/Qualifiers
SOURCE	1. 119914
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/db_xref="taxon:3702"
	/chromosome="1"
	/clone="F21M1"
	/ecotype="Columbia"
misc_feature	1. 9262
	/note="overlap with bases 68998-78259 of 'IGF' BAC clone F21M1."
misc_feature	1. 7000
	/note="Proteins in this region are annotated in the F21M1 entry, AC002411."
gene	7684. 10913
	/gene="F21M1.1"
CDS	join(7684..7866,7965..8027,8122..8160,8250..8361,8458..8728,8875..9893,10003..10256,10344..10410,

10501..10551,10685..10781,10885..10913)  
/gene="F21M11.1"  
/note="Hypothetical protein"  
/codon start=1  
/db\_xref="GI:4204284"  
/protein\_id="AA010665.1"  
/translation="MAPNLRIKACDAMKLLIGSEKTPAFIRKLKTYENMDPIEE  
DAYKVLDAIFDEADQSTENKKEEKKKEEKSRSVATSRGRKAPELVODE  
DMDEDEPLKRLRSRGRASSSSSSSYNNEDLQTPBEDDDQVTEPLKRY  
VRRNGRGLAMTYNNASPSSSRLSMEPEEYPMVLLPAHMETKVSASALVIND  
ENIDHKPVISDTCNSAPMLEWGNHVOEDMETKXLDLNTDTAMDVSPSALGES  
SEKVAASAVELASSTSGAKICLSAPATGERTNHLPSMEDIPRAMEKCKSKYKI  
VHEPESVLGPMKDCSCYIDLAKNSTQLTETVCDMSKAGDESGAVISNPLVVP  
ECEISGDGKAIASNMKDITAGEENVEIPWNEINENPSPFRMPSPFVQDAVPLVS  
LSFSDQSCSTSCIEDCLASWSCNCAIGVDPAYTLTDLGLKEEFELARISEADQ  
RVOVLFCSECPLEPAKVEILEPCGHLKRGAIKECFKCGCTKRGVVRQGMN  
KLOPPTPGKMGGLATLEKLPKGAFCIGYGLILIPELYORSPEDKTLPVYLDAAH  
WGSBERLEDKALCLDGMFYGNISRLNRCDAANIETPVQETPDQHYHLAFPTT  
RDIEMAEELAMDYGDIFNDNDSLMPDCLCGSRFCRNKRSKTKMQILNKA"  
/gene="F21M11.2"  
/join(13033..13488,13717..13926,14139..14514)  
/gene="F21M11.2"  
/join(13061..13488,13717..13926,14139..14316)  
/gene="F21M11.2"  
/note="Similar to acid phosphatase; location of ESTs  
110C2Y7, gb|T42036, and 110C2XP, gb|A1100245"  
/codon start=1  
/protein\_id="AA010666.1"  
/db\_xref="GI:4204285"  
/translation="MDRTWFLSLITIASLVGVVSAGDMNINLQRLGSSSSONGIVS  
KGITDLKGVCESWRINVEYHNRKPDVVOECVSHIKDYMSSQYKDVAVDVSI  
LHSGSCSKSCGDMDAMFIDIDTLSTIPIYKHKGFEGGCEKADFEWDIOKK  
APAVPMKLIYHDIRERGIKIFLISRRKEATVDNLIOAGYCGMSMLMRGLDQ  
QKEVKYKSEKRWIMSLGYRWVGWDQSSPAGCPLPRRTFKLNSIYYVA"  
/complement(14004..16362)  
/gene="F21M11.3"  
/complement(join(14004..14312,14565..14641,14787..15831,  
15906..16006,16327..16362))  
/gene="F21M11.3"  
/complement(join(14567..14641,14787..15831,15906..16006,  
16327..16362))  
/gene="F21M11.3"  
/note="Unknown protein; location of ESTs 40C3T7,  
gb|AA728590 and40C3T7, gb|T04573"  
/codon start=1  
/protein\_id="AA010667.1"  
/db\_xref="GI:4204286"  
/translation="MLDQISGNHTHEKLSVETAEPPHLLNDRHVIEIRKASVPIPT  
ICDEAEKCSPTYSRKYVDSKRYVEHVEEVSLEENKEEVSERKSLKSK  
TDQDILEVANSNGSYENHRYKNCRESDDIDEDEDCSDLDDEEESYDVGSE  
DSLHPNTEKYVTDIDGKTEIDSKLRSNETHYRDEGGLANVENLITQWKA  
SKRTKOKQOKENSPIADQEKRDSSSGTPOIDITLSIKPKRRIEKKLRQ  
LAVDASLSTLSTSESGSCNSASMTLPEKLSKSYCSKPIRINDDPVCALTL  
EDIKQFATSTPRKSPKSPDETPIIGTVGSGWNRKALICGSASSFKGIPMTSY  
REKSVVHSTPEARLEKALNNIDK"  
/complement(16835..20238)  
/gene="F21M11.4"  
/complement(join(16835..17185,17274..17392,17491..17656,  
17795..17885,17982..18079,18175..18361,18504..18604,  
18705..19049,19134..19349,19439..19791,19862..19967,  
20188..20238))  
/gene="F21M11.4"  
/note="Hypothetical protein"  
/codon start=1  
/protein\_id="AA010669.1"  
/db\_xref="GI:4204288"  
/translation="MGSSSPEARARACVPSMILIFLEICTVHYTNRRKLNDAVISA  
NLNIPKTRRDLRFMESVLSIYKSLAAVSVLPOQIIPNDQYKNDALNNSNPH  
GESDESEMDKVSRCGSDPSRDLPTSESDPRPDQMTKQSLDLYE  
FSESDYDANHTPESYTEOAKNVRDITASOPSAAKRI CGSFIQESSPNRTO  
TLRLMESLSDPTDYKQONQMVYFSFWIYLSFVLCARSNNLRVVVFCVQOQJPK  
SHTEDSKRKRDITASDAMENHLKVRRENMLQKSDADICNGKCSANSDQJSEKIS

KALBOTSNITTCGFCOSARVSEATGEMLYHSGRPVVDGDIIFRSNVIHVSACIEMA  
POVYREGTVTKRLKELAKMKIKCTKSLGGAALGCVKSGRBYHPCAREISCR  
WDYDFLLCPHSSVKTPEKSGRVSRARPELPIKINAELESLEQTAFTPELVLCG  
SALSXDKKLMESLAVRFNATISRWNSVTHVASDTEKACTRLKMLKMLGK  
IINAAVMKASLAKASQVDEBPPEIQIDQGCODGKTRARLRATETKPKLFEELKFF  
GDFYKVEDLQNLVKVAGGTILNTEDEGASSNNVNDORSISIVVNIIDPHFCAL  
GEEVTIIMORANDALASQSGSRVGHVTVLESJAGYKHLHVIG"  
20752..24647  
/gene="F21M11.5"  
/join(20752..20994,21362..21497,21596..21740,21825..21994,  
22102..22118,22542..22758,22920..23064,23343..23442,  
23599..23693,24021..24072,24227..24298,24441..24647)  
/gene="F21M11.5"  
/note="Hypothetical protein"  
/codon start=1  
/protein\_id="AA010668.1"  
/db\_xref="GI:4204287"  
/translation="MGANSKSVTASFVIAVFLICGRTAVEDETEPHDYSKLSGI  
IIPGRASQDLRAWSILDCPYPLDNPDLVWLDTTKLSAVNCPKCVLPVNDQD  
HPECKSPDSGLSATTEDDPGITTPPLSTVWKEMKCYBERIENATVAVPYDMRLS  
PYKLEBRDLVFHKLKLTETLAKLNGRPVIFAHSMGNVFRYFLEMLLETAPKYL  
KMLDQIHAFVAVGAPLGSVBAIKSTLSGTFGLPVEGTARLNSPSSLSLMLPF  
SKNGKDNATFMTFSGAAKDKRKYVHCDEEYQSGMPTNIIINIEIPSTARBELA  
DGLTKAIEDVDYDPSKMLHOLKTVFPFVNIHRSLSAGFLYHDDPVNPLTPW  
RRPPIKNVPCIGAHLKTEVGVYFAPSGKPYDNNIITDIYETEGSLVSRGTVDG  
NAGPITGDETVYHSLSMCKMNLGKRNITNAPQILIKIKQPEHSDVHVELND  
HHGSDIILANMTKARPVKITTYEDESIPGRITVWELDKSGY"  
27777..28734  
/gene="F21M11.6"  
/db\_xref="GI:4204289"  
/translation="NMTKTRLPFRPVLTADRKERDAFISVTDNPPETARFPP  
KLVPVNPDISKSSNTAAAEPIGSNQMLAGLSHLYTQGLPFGQMQVDAQESS  
KIKPSHTVEPAECEPKRKYREVNMLRSDDAOLPGIVNPAOLARFLKL"  
/complement(29264..32033)  
/gene="F21M11.7"  
/complement(join(29264..31015,31312..31414,31484..31587,  
31687..31787,31887..31987,32087..32187,32287..32387,32487..32587,  
32687..32787,32887..32987,33087..33187,33287..33387,33487..33587,  
33687..33787,33887..33987,34087..34187,34287..34387,34487..34587,  
34687..34787,34887..34987,35087..35187,35287..35387,35487..35587,  
35687..35787,35887..35987,36087..36187,36287..36387,36487..36587,  
36687..36787,36887..36987,37087..37187,37287..37387,37487..37587,  
37687..37787,37887..37987,38087..38187,38287..38387,38487..38587,  
38687..38787,38887..38987,39087..39187,39287..39387,39487..39587,  
39687..39787,39887..39987,40087..40187,40287..40387,40487..40587,  
40687..40787,40887..40987,41087..41187,41287..41387,41487..41587,  
41687..41787,41887..41987,42087..42187,42287..42387,42487..42587,  
42687..42787,42887..42987,43087..43187,43287..43387,43487..43587,  
43687..43787,43887..43987,44087..44187,44287..44387,44487..44587,  
44687..44787,44887..44987,45087..45187,45287..45387,45487..45587,  
45687..45787,45887..45987,46087..46187,46287..46387,46487..46587,  
46687..46787,46887..46987,47087..47187,47287..47387,47487..47587,  
47687..47787,47887..47987,48087..48187,48287..48387,48487..48587,  
48687..48787,48887..48987,49087..49187,49287..49387,49487..49587,  
49687..49787,49887..49987,50087..50187,50287..50387,50487..50587,  
50687..50787,50887..50987,51087..51187,51287..51387,51487..51587,  
51687..51787,51887..51987,52087..52187,52287..52387,52487..52587,  
52687..52787,52887..52987,53087..53187,53287..53387,53487..53587,  
53687..53787,53887..53987,54087..54187,54287..54387,54487..54587,  
54687..54787,54887..54987,55087..55187,55287..55387,55487..55587,  
55687..55787,55887..55987,56087..56187,56287..56387,56487..56587,  
56687..56787,56887..56987,57087..57187,57287..57387,57487..57587,  
57687..57787,57887..57987,58087..58187,58287..58387,58487..58587,  
58687..58787,58887..58987,59087..59187,59287..59387,59487..59587,  
59687..59787,59887..59987,60087..60187,60287..60387,60487..60587,  
60687..60787,60887..60987,61087..61187,61287..61387,61487..61587,  
61687..61787,61887..61987,62087..62187,62287..62387,62487..62587,  
62687..62787,62887..62987,63087..63187,63287..63387,63487..63587,  
63687..63787,63887..63987,64087..64187,64287..64387,64487..64587,  
64687..64787,64887..64987,65087..65187,65287..65387,65487..65587,  
65687..65787,65887..65987,66087..66187,66287..66387,66487..66587,  
66687..66787,66887..66987,67087..67187,67287..67387,67487..67587,  
67687..67787,67887..67987,68087..68187,68287..68387,68487..68587,  
68687..68787,68887..68987,69087..69187,69287..69387,69487..69587,  
69687..69787,69887..69987,70087..70187,70287..70387,70487..70587,  
70687..70787,70887..70987,71087..71187,71287..71387,71487..71587,  
71687..71787,71887..71987,72087..72187,72287..72387,72487..72587,  
72687..72787,72887..72987,73087..73187,73287..73387,73487..73587,  
73687..73787,73887..73987,74087..74187,74287..74387,74487..74587,  
74687..74787,74887..74987,75087..75187,75287..75387,75487..75587,  
75687..75787,75887..75987,76087..76187,76287..76387,76487..76587,  
76687..76787,76887..76987,77087..77187,77287..77387,77487..77587,  
77687..77787,77887..77987,78087..78187,78287..78387,78487..78587,  
78687..78787,78887..78987,79087..79187,79287..79387,79487..79587,  
79687..79787,79887..79987,80087..80187,80287..80387,80487..80587,  
80687..80787,80887..80987,81087..81187,81287..81387,81487..81587,  
81687..81787,81887..81987,82087..82187,82287..82387,82487..82587,  
82687..82787,82887..82987,83087..83187,83287..83387,83487..83587,  
83687..83787,83887..83987,84087..84187,84287..84387,84487..84587,  
84687..84787,84887..84987,85087..85187,85287..85387,85487..85587,  
85687..85787,85887..85987,86087..86187,86287..86387,86487..86587,  
86687..86787,86887..86987,87087..87187,87287..87387,87487..87587,  
87687..87787,87887..87987,88087..88187,88287..88387,88487..88587,  
88687..88787,88887..88987,89087..89187,89287..89387,89487..89587,  
89687..89787,89887..89987,90087..90187,90287..90387,90487..90587,  
90687..90787,90887..90987,91087..91187,91287..91387,91487..91587,  
91687..91787,91887..91987,92087..92187,92287..92387,92487..92587,  
92687..92787,92887..92987,93087..93187,93287..93387,93487..93587,  
93687..93787,93887..93987,94087..94187,94287..94387,94487..94587,  
94687..94787,94887..94987,95087..95187,95287..95387,95487..95587,  
95687..95787,95887..95987,96087..96187,96287..96387,96487..96587,  
96687..96787,96887..96987,97087..97187,97287..97387,97487..97587,  
97687..97787,97887..97987,98087..98187,98287..98387,98487..98587,  
98687..98787,98887..98987,99087..99187,99287..99387,99487..99587,  
99687..99787,99887..99987,100087..100187,100287..100387,100487..100587,  
100687..100787,100887..100987,101087..101187,101287..101387,101487..101587,  
101687..101787,101887..101987,102087..102187,102287..102387,102487..102587,  
102687..102787,102887..102987,103087..103187,103287..103387,103487..103587,  
103687..103787,103887..103987,104087..104187,104287..104387,104487..104587,  
104687..104787,104887..104987,105087..105187,105287..105387,105487..105587,  
105687..105787,105887..105987,106087..106187,106287..106387,106487..106587,  
106687..106787,106887..106987,107087..107187,107287..107387,107487..107587,  
107687..107787,107887..107987,108087..108187,108287..108387,108487..108587,  
108687..108787,108887..108987,109087..109187,109287..109387,109487..109587,  
109687..109787,109887..109987,110087..110187,110287..110387,110487..110587,  
110687..110787,110887..110987,111087..111187,111287..111387,111487..111587,  
111687..111787,111887..111987,112087..112187,112287..112387,112487..112587,  
112687..112787,112887..112987,113087..113187,113287..113387,113487..113587,  
113687..113787,113887..113987,114087..114187,114287..114387,114487..114587,  
114687..114787,114887..114987,115087..115187,115287..115387,115487..115587,  
115687..115787,115887..115987,116087..116187,116287..116387,116487..116587,  
116687..116787,116887..116987,117087..117187,117287..117387,117487..117587,  
117687..117787,117887..117987,118087..118187,118287..118387,118487..118587,  
118687..118787,118887..118987,119087..119187,119287..119387,119487..119587,  
119687..119787,119887..119987,120087..120187,120287..120387,120487..120587,  
120687..120787,120887..120987,121087..121187,121287..121387,121487..121587,  
121687..121787,121887..121987,122087..122187,122287..122387,122487..122587,  
122687..122787,122887..122987,123087..123187,123287..123387,123487..123587,  
123687..123787,123887..123987,124087..124187,124287..124387,124487..124587,  
124687..124787,124887..124987,125087..125187,125287..125387,125487..125587,  
125687..125787,125887..125987,126087..126187,126287..126387,126487..126587,  
126687..126787,126887..126987,127087..127187,127287..127387,127487..127587,  
127687..127787,127887..127987,128087..128187,128287..128387,128487..128587,  
128687..128787,128887..128987,129087..129187,129287..129387,129487..129587,  
129687..129787,129887..129987,130087..130187,130287..130387,130487..130587,  
130687..130787,130887..130987,131087..131187,131287..131387,131487..131587,  
131687..131787,131887..131987,132087..132187,132287..132387,132487..132587,  
132687..132787,132887..132987,133087..133187,133287..133387,133487..133587,  
133687..133787,133887..133987,134087..134187,134287..134387,134487..134587,  
134687..134787,134887..134987,135087..135187,135287..135387,135487..135587,  
135687..135787,135887..135987,136087..136187,136287..136387,136487..136587,  
136687..136787,136887..136987,137087..137187,137287..137387,137487..137587,  
137687..137787,137887..137987,138087..138187,138287..138387,138487..138587,  
138687..138787,138887..138987,139087..139187,139287..139387,139487..139587,  
139687..139787,139887..139987,140087..140187,140287..140387,140487..140587,  
140687..140787,140887..140987,141087..141187,141287..141387,141487..141587,  
141687..141787,141887..141987,142087..142187,142287..142387,142487..142587,  
142687..142787,142887..142987,143087..143187,143287..143387,143487..143587,  
143687..143787,143887..143987,144087..144187,144287..144387,144487..144587,  
144687..144787,144887..144987,145087..145187,145287..145387,145487..145587,  
145687..145787,145887..145987,146087..146187,146287..146387,146487..146587,  
146687..146787,146887..146987,147087..147187,147287..147387,147487..147587,  
147687..147787,147887..147987,148087..148187,148287..148387,148487..148587,  
148687..148787,148887..148987,149087..149187,149287..149387,149487..149587,  
149687..149787,149887..149987,150087..150187,150287..150387,150487..150587,  
150687..150787,150887..150987,151087..151187,151287..151387,151487..151587,  
151687..151787,151887..151987,152087..152187,152287..152387,152487..152587,  
152687..152787,152887..152987,153087..153187,153287..153387,153487..153587,  
153687..153787,153887..153987,154087..154187,154287..154387,154487..154587,  
154687..154787,154887..154987,155087..155187,155287..155387,155487..155587,  
155687..155787,155887..155987,156087..156187,156287..156387,156487..156587,  
156687..156787,156887..156987,157087..157187,157287..157387,157487..157587,  
157687..157787,157887..157987,158087..158187,158287..158387,158487..158587,  
158687..158787,158887..158987,159087..159187,159287..159387,159487..159587,  
159687..159787,159887..159987,160087..160187,160287..160387,160487..160587,  
160687..160787,160887..160987,161087..161187,161287..161387,161487..161587,  
161687..161787,161887..161987,162087..162187,162287..162387,162487..162587,  
162687..162787,162887..162987,163087..163187,163287..163387,163487..163587,  
163687..163787,163887..163987,164087..164187,164287..164387,164487..164587,  
164687..164787,164887..164987,165087..165187,165287..165387,165487..165587,  
165687..165787,165887..165987,166087..166187,166287..166387,166487..166587,  
166687..166787,166887..166987,167087..167187,167287..167387,167487..167587,  
167687..167787,167887..167987,168087..168187,168287..168387,168487..168587,  
168687..168787,168887..168987,169087..169187,169287..169387,169487..169587,  
169687..169787,169887..169987,170087..170187,170287..170387,170487..170587,  
170687..170787,170887..170987,171087..171187,171287..171387,171487..1

Db 20992 AAGTCGCGATCTTCATTTCTCGCTGCTTATTTCTGCGGCGAGTCACCTTGTTGATG 21051  
 QY 81 ----- 81  
 Db 21052 AATTCGAAGCAAAATATAGCAATGAAGCATGTCGTCTCTCTATTTGATGCTTCATTA 21111  
 QY 81 ----- 81  
 Db 21112 GTCAACAGTCAGCGCTTCTGAATCTGAGTTTAGAGTCATATAAAACAGCTGACTCGCGGAG 21171  
 QY 81 ----- 81  
 Db 21172 TGTTCOCATCGCTTTGTTGGCTAAATGTAGCGCAATATGTATATTAAGTCGCGC 21231  
 QY 81 ----- 81  
 Db 21232 TTTTATTCACATAGATCTGCAAGTTTTCAGAGTGTCAATAGTAGTAAAGAAATGTTA 21291  
 QY 81 ----- 81  
 Db 21292 GGTCAATTTACTTGTCAGTTCGATTCCTTTGGTTGCTTACTGATCGAGTCGATGGA 21351  
 QY 82 -----LeuLeuSerAlaValAsnCySTrpPheLysCySmeValLeuAspProTy 98  
 Db 21352 TGGTTTACAGCTTCTTCTGCTGTCAACTGCTGTTTAACTGATGCTGCTAGATCTTA 21411  
 QY 98 rAnGlnThrAspHisPProGluCyAluSeraTrpAspSeraGlyLeuSerAlaLeuH 118  
 Db 21412 TATATCAACAGACCAATCCGAGGTGAATGACGCGCTGACAGTGTCTTTCAGCCATCAC 21471  
 QY 118 rGluLeuAspProGlyTyTrIleThr----- 126  
 Db 21472 AGAATTGGATCCAGTTACATACACAGTAGTTTCGATTTTCTTCTTTCAGCTTTC 21530  
 QY 126 ----- 126  
 Db 21531 TTCAATTGATATCATCTTGTGTGATATATATGCTAAGTTCATTAATTTGCTCAAT 21590  
 QY 127 ----GlyProLeuSerThrValTTrpLysGluTrpLeuLysTrpCyValGluPheGly 145  
 Db 21591 TTCAGGTCTCTTCTTACTGTCTGAAAGAGTGGCTTAAGTGTGTGAGTTGGTAT 21650  
 QY 145 eGluAlaAsnAlaIleValAlaValProTyAspTrpArgLeuSerProThrLysLeuG 165  
 Db 21651 AGAAGCAATGCAATTCGTGCTGCTTCATACGATTTGAGATTTTCACCAACCAATTGA 21710  
 QY 165 uGluAspLeuTyTrPheHisLysLeuLys----- 175  
 Db 21711 AAGGCGTGAACCTTACTTCACAAAGCTCAAGTTAGTCTTATCAGGCTAATGCTTTTAT 21770  
 QY 176 -----LeuH 177  
 Db 21771 CTTCTCTTTTATGTAAGATTAAGCTAAGAGCTGCTGCTCTTCTTTTTCAGGCTGAC 21830  
 QY 177 rPheGluThrAlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHisSerMetG 197  
 Db 21831 CTTTGAACCTGCTTAAACTCCGTGCGGCTCTTCTATAGTATTTGCCCATTCATGCG 21890  
 QY 197 yAsnAsnValPheArgTyTrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisTy 217  
 Db 21891 TATATATGCTCTCAGATACTTCTCGAATGCTGAGGCTGAAATTCACAAACAAATTA 21950  
 QY 217 rLeuLysTrpLeuAspGlnHisIleHisAlaIleTyTrPheAlaVal----- 231  
 Db 21951 TTTGAAGTGGCTTGATCAGATATCATGCTTATTTTCGCTGT- TGGTACCGGCTTACTAT 22009  
 QY 231 ----- 231  
 Db 22010 CCTTAAGTACCATTTTATTTTCTCTAATTTGGGGAGTATGTTGTACTTACGAT 22069  
 QY 232 -----GlyAlaProLeuLeuGlySeraGluAl 241  
 Db 22070 TGAGCTGCATACCTGATTTGTTGTTGATTTAGAGAGCTCTCTTCTTGTGTTGAGGC 22129

QY 241 aIleLysSerThrLeuSerGlyValIlePheGlyLeuProValSeraGlu----- 257  
 Db 22130 AATCAAACTACTCTCTCTGGTGTACGTTTGGCTTCCTCTTCTGTAGAGTACCTTGA 22189  
 QY 257 ----- 257  
 Db 22190 CTTCTTTAGTTTAAAGTAGTATATCAACAGGCTTAAATCACTGAGTTTTCCT 22249  
 QY 257 ----- 257  
 Db 22250 TTTGAAGTATTAATTTTGTAAATGAACCTGCTGACGATATGATCTGTAGATCTT 22309  
 QY 257 ----- 257  
 Db 22310 GAAGTCTAGTTATCAAAAGAACATATTTGGGTAGTATACCTGTCAGCGGCTTACTAA 22369  
 QY 257 ----- 257  
 Db 22370 TACAACCAACACATGTACACTGATTTAGTTTCAGATTTATATGTAGACTTTAAGTT 22429  
 QY 257 ----- 257  
 Db 22430 GAGAAAGAACTTGTGACTGAAATCTTTTATTTTAAAGCTATGATTTGTTATTAAT 22489  
 QY 258 -----GlyThrAl 260  
 Db 22490 CATGTGACATATATGACATGCGCTTCTCATGTTTTTGTGGCAAGCTTCAGAGAACTGC 22549  
 QY 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCy 280  
 Db 22550 TCGGTTGTTGTCATATCTTTTGGCGTGCATATGTTGGCTTATGACATTTTCAAGAAATTG 22609  
 QY 280 eLysGlyAspAsnThrSerTrpThrHisPheSerGlyGlyValAlaLysLysAspLysAr 300  
 Db 22610 CAAGGCTATATACATTTCTGACGCAATTTTCTGGGGTCTCTGAAAGAAAGATTAACG 22669  
 QY 300 gValTyTrHisCyAspArgLugLugLugLysTyTrGlnSerLysTyTrSerGlyTrpProThrAsn 320  
 Db 22670 CGTATACCACTGTGATGAAGAGAAATATCAATCAAAATATTTCTGCTGCGCACAAATAT 22729  
 QY 320 eIleAsnIleGluIleProSerThrSer----- 329  
 Db 22730 TATTAACATTTGAATTTCTTCCACTAG- CGGTTAGACTGTATATGCAACTGTAACT 22788  
 QY 329 ----- 329  
 Db 22789 AACAAAGTTTCACCAAGAAATGTTCACTGCATATTTGCTTGTGATGTGATCCATC 22848  
 QY 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCySglLysProThrLeu 349  
 Db 22849 AGTTACAGAAACAGCTTATGTCACATGACACAGCATGGAATGTGGCTTCCACACCTTTT 22908  
 QY 349 uSerPheThrAlaArgLugLugLugLysAspGlyTyThrLeuPheLysAlaIleGluAspTyTrAs 369  
 Db 22909 GCTTTTACAGCGCGCTGAACCTAGATGGGACTCTTTTCAAGCAATAGAAAGACTATGA 22968  
 QY 369 pProAspSerLysAlaGlyMetLysHisGlnLeuLysLysLeu----- 382  
 Db 22969 CCCAGATGACAGAGAGATGTTACACACAGTTAAAGAAATGA- CGTACCTTTCTTTGTGATTA 23027  
 QY 382 ----- 382  
 Db 23028 GAAATATTTGCTATGATCATCATCTTGCCTTCTTGTAGCTCAAAATTTTGTGTTAA 23087  
 QY 382 ----- 382  
 Db 23088 ATCTCATATCAATTTGTCATATGCTTGTCTTCTTACTATTAAGAAACAAAGTATATCA 23147  
 QY 382 ----- 382  
 Db 23148 GAAACCTTATTAATTAATATCAGTTCTCTCTTATATTAATGAATGCTTTTTCGTTTAC 23207

QY 382 ----- 382  
 Db 23208 AGTTATGAATGCAAAAGGGGATTTTGTAGTGTGATTCCTCATCTCTAGTTGTTT 23267  
 QY 382 ----- 382  
 Db 23268 TGACATAAGCGCAATTTTGTCTTTCAGCAAACTCTTGGAATTATATATAACATGCT 23327  
 QY 383 ----- 396  
 Db 23328 AACTATACCTTTTCAGGTTGTATCATGATGACCTGTTTTTAATCCTCTGACTCTTGGA 23387  
 QY 396 ----- 414  
 Db 23388 GAGACACACCTATAAAAAATGTATTTGCTATATGCTCATCTAAAGACAGAGGTATG 23447  
 QY 414 ----- 414  
 Db 23448 ATGCATTTCTCAATATCATATTATGCTTGCCTTGTATATATCCCATTTGGTTGC 23507  
 QY 414 ----- 414  
 Db 23508 AATATCTTTTGAATTATGATTTATCTCTCCCTTGCACTCTATGCTATTAAGGTTAA 23567  
 QY 415 ----- 424  
 Db 23568 GGTACTAATATGATGAAGCTGCTGTCATAGTTGTTTACTTTGCCCCAAGTGCGCA 23627  
 QY 424 ----- 444  
 Db 23628 ACCTTTTCCGATTAATGATGATCAGGATATCATTTATGAAGAGTTCCCTCGT 23687  
 QY 444 ----- 445  
 Db 23688 GTCAAGGTAATTTTCCGCAATGCGAAGAAAGTAAAGGAAAGCAAGTCTTCTGTATCAG 23747  
 QY 445 ----- 445  
 Db 23748 TCTAGTGGCATGTTATCTCAGTTGTCATAGCAAAATTTTAAACAATAATTAAAGTAC 23807  
 QY 445 ----- 445  
 Db 23808 TTTTATCATCTCTTTGAGCTTGTAGATGATGATCAGTGGCTTAAAGTGGAAGAGTGT 23867  
 QY 445 ----- 445  
 Db 23868 TGCATGAACAATGACACTTGTATCAAAAGATTAACGAAAAACAATAACCATTTCTG 23927  
 QY 445 ----- 445  
 Db 23928 AATTCATATTATTAGAGTAGTGTGCTTTTAAAAAATTTGTTTAAAGAAACCGAAAA 23987  
 QY 446 ----- 455  
 Db 23988 CTAGTTCATATCTTGTATGTCATATCTGAGGTCGGAACGTGTGTATGAGGAAAGC 24047  
 QY 455 ----- 463  
 Db 24048 TGAACCTATATAGGGGATGAGACGCTAGCTCAGAAAGTTGTTGAAATTAATCTTCTT 24107  
 QY 463 ----- 463  
 Db 24108 GCAAACTAGCAAGACTAAGATTAATTAATCTGCTTGTGAAACACTGCTGTATGTTCTTA 24167  
 QY 464 ----- 464  
 Db 24168 GTACACTGCAATATGACTCCGCTACTTTTATGATTAATGAATGATCTTATAGG 24227  
 QY 464 ----- 484  
 Db 24228 TACCTTATCATCTCTCTTGGTGGAAGATTGGCTCGGACCTTAAGTTAACAATAACA 24287  
 QY 484 ----- 486  
 Db 24287 etala-Pro----- 486

Db 24288 TGGCTCCCAAGGACTCTTTTGTAGTCTCTCACTTATATAGATCAAACTTAAGTAC 24347  
 QY 486 ----- 486  
 Db 24348 TTTTCTGTTATGTGTATTAACCTCAATTTGTTCTTTCTAAAAATCATATATCTCTG 24407  
 QY 487 ----- 487  
 Db 24408 TACTCTCAAGAACTGTATTAATCTTAAACGAGATTCATTTGGCAAAATTAACAACAG 24467  
 QY 488 ----- 507  
 Db 24468 CCAAGACACAGATGAGACGACCTGATCTGGAATCTTAAGTTGATCATGACATGGGTCA 24527  
 QY 508 ----- 527  
 Db 24528 GACATCATAGCTTAACATGACCAAAAGCACCAGGTTAAGTACATAACCTTTTATGAAGAC 24587  
 QY 528 ----- 546  
 Db 24588 TCTGAGAGGATTCGGGGAAGAACCGACGCTGGAGCTTGATTAAGTGGTAT 24644  
 RESULT 8  
 AF493159 2609 bp mRNA linear PLN 22-SEP-2004  
 LOCUS Medicago truncatula lecithine cholesterol acyltransferase-like  
 DEFINITION protein mRNA, complete cds.  
 ACCESSION AF493159  
 VERSION AF493159.1 GI:25992000  
 KEYWORDS  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;  
 Medicago.  
 REFERENCE 1 (bases 1 to 2609)  
 AUTHORS Noiret,A., Benveniste,P., Banas,A., Styenne,S. and Bouvier-Nave,P.  
 TITLES Expression in Yeast of a novel phospholipase A1 cDNA from Arabidopsis thaliana Eur. J. Biochem. 271 (18), 3752-3764 (2004)  
 JOURNAL 1535352  
 PUBMED 2 (bases 1 to 2609)  
 REFERENCE Benveniste,P., Bouvier-Nave,P., Schaller,H. and Noiret,A.  
 AUTHORS Direct Submission  
 TITLES Submitted (15-MAR-2002) Plant Molecular Biology Institute, CNRS, 28 rue Goethe, Strasbourg 67083, France  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1..2609  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone\_lib="Samuel Roberts Noble Foundation Medicago truncatula insect herbivory library (Korth, K., et al., unpublished)"  
 /note="fabaceae; sequence derived from EST clones BE321377 and B1267156"  
 82..1980  
 /note="enzyme involved in the metabolism of phospholipids and sterols"  
 /codon\_start=1  
 /product="lecithine cholesterol acyltransferase-like protein"  
 /protein\_id="AAU77002.1"  
 /db\_xref="GI:25992001"  
 /translation="MANKPRLISIIILAPVAVGSGSGSELDYSKLSGIIRGFAS  
 TQIRAWSILDCPSLDPNPLDVLMDITTKLSAVNCMLCLDLPNQTDRHDPCKSR  
 PDGSLGITELELDGYITGPLSSVWKMIKCFEIGAVNIIVPYDWRSPSLER  
 DLVFKLKLFEPAFLRGGSILVFGHSLGNVRFLEWLKLEIAPKHVIOWDQHI  
 HAVFAVALPLGATEITLSEFTGLPVSEGTALIMFSPASSIMWPFYSKCRAS  
 NKYMKHPSGGKQVGTNYHGDSESPFNFGWPKLIIINIRIPSTRGEAVPSSEIPE  
 ANLSGMECGIPLTQLSFSAKRIADGSFFKAIEDDPDSKRLLYLLEKSYLDPLVLPNT





Qy 519 VAllyVtYrIleThPheTyrGluApSergIuSerIlleProGlyLyLeaRgThraIaVal 538  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :  
Db 1630 GTCAAGTACATACCACTATTATGAGATCTGAAAGTATCCAGGAAGAACAGCTGTT 1689  
Qy 539 TTPGluLeuApSergIyTyr 546  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :  
Db 1690 TGGAGCTTGATTAACCAATCAC 1713

RESULT 10  
AP008208\_348/c  
WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

Fragment Name	Begin	End
AP008208_000	1	110000
AP008208_001	100001	210000
AP008208_002	200001	310000
AP008208_003	300001	410000
AP008208_004	400001	510000
AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000
AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000
AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000

AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	10110000
AP008208_101	10100001	10210000
AP008208_102	10200001	10310000
AP008208_103	10300001	10410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000





Db 105862 TGGAAAAAACCATGCGTGGCAAGTGAAGATGATAAAATTGACATCAAGCTTTTGT 105803  
Qy 126 ----- 126  
Db 105802 GAGTTAATGATGATATTGCCATGAGAAATATAGCATTAGCTAGTTACATTGCCAAATTTC 105743  
Qy 126 ----- 126  
Db 105742 ATGTTCAATCGGTAGCTGTGAATTCCTCCATCTCGTTCTCTCAAGAAATAAAAAATGATT 105683  
Qy 127 ----- GlyProle 129  
Db 105682 ACCCATAAAAAAATCTTTCTTACTAATTTGGTTTATTGGCCAAAGTTTTCAGGTCTCT 105623  
Qy 129 uSerThrVal1TTPLySGlUTTPLeuLySTpCySVal1GluPheGly11Glu1AlaAsn1 149  
Db 105622 TTCTTCAGTGTGGAAGAAATGGGTCAAAATGCTGTGAAGAAATTTGGTATGGAAGCTAATTC 105563  
Qy 149 Al1eVal1AlaVal1ProLyAsp1TpaRbLeuSerProThrLybLeuGluGluArgAsp1e 169  
Db 105562 CATTAATCGGGGTTCCATACGATTTGGAGGCTGCCCTCCCTCGATGCTTGAGAGAGGGATCT 105503  
Qy 169 uTyRPhen1sLybLeuLyS ----- 175  
Db 105502 GTAATCCATAAATTAAAGTTTGTAACACTTCATTCATATACATAGCAATCAATTTAC 105443  
Qy 175 ----- 175  
Db 105442 TCACCTGCCATTGGCTTTCGATAGTATTGATGTTATTACTAATAATGGTGTAAAAACA 105383  
Qy 175 ----- 175  
Db 105382 CTTCTAAATTGTRATCAACAGAAATATACATCTTGTTTACCTTCATACCTCAAAAAA 105323  
Qy 175 ----- 175  
Db 105322 TCACGATAACTAAGTACTAGTATCAGAAATACAAAAGTTGGTGCCTTGGTG 105263  
Qy 175 ----- 175  
Db 105262 CATGCAAGCTGAATGATTAATTTTTTGGGGGGAATGATAAGCTTTTACTTGCCT 105203  
Qy 175 ----- 175  
Db 105202 CTTTTCATTTTTTATGACATCTTTCACCTGAATTAAGTTGTTTCTTCAAGTTTA 105143  
Qy 175 ----- 175  
Db 105142 TACAATACTGTAATATGACACGGCGATATGTCGGAATTCATTGTTGATGCGGAAGAT 105083  
Qy 176 ----- leuThrPheGluThrAlaLeuLyL 184  
Db 105082 AGTATTTGGTGCTCATAGTCCCTTTAACTGACGGTTACTTTTGAACGTCATTAAGAAC 105023  
Qy 184 euArgLyGlyProSer11eVal1PheAlaHisSerMetGlyAenAnVal1PheArgTyR 204  
Db 105022 TTTCGAGAGAGGCTTCTTATAGTGTTCATTCATTCGATTAATAGTGTTCGCTACT 104963  
Qy 204 heLeuGluTTPLeuArgLeuGlu11eAlaProLySh1sTyR1eulysTTPLeuAspGlnH 224  
Db 104962 TTCTCGAATGGTTGAACCTGAAGATCGCTCCCAAGCATTCATCGATGGCTTGACGAAC 104903  
Qy 224 Is11eHisAlaTyRPhen1AlaVal1 ----- 231  
Db 104902 ATATACATGATACCTTTCAGTGTGATGTAAACAGAAAGAAATTAATTCATCACCATC 104843  
Qy 232 ----- GlyAlaProLeuLeuGlySerV 239  
Db 104842 GTTATTTGTTAGTGCATGCCATTTTTCCTTGCTTTCAGGTGACCTCTTCTTGTTCTTA 104783  
Qy 239 a1GluAla1eLySserThrLeuSerGlyVal1ThrPheGlyLeuProVal1SerGlu --- 257  
Db 104782 CTGAACACAGTTAAAGCTGCTCTTTCGAGCAACATTTGGTCTTCAGTCCAGCGAGGTCA 104723

Qy 257 ----- 257  
Db 104722 GCCTTATGAATTTATATATTTTTCATATATGTGTCCTTGCCCTCAGAGGCTATATTTC 104663  
Qy 257 ----- 257  
Db 104662 TGATATTATACATTTTCTCCAAACATTTGATGTTGAATTAAGTATGTTGGTT 104603  
Qy 257 ----- 257  
Db 104602 GTCCAAATACGTAGATTATTTATTTCTTTTGGATTGCACTAGATTTTGTGTGATTCTT 104543  
Qy 257 ----- 257  
Db 104542 TTGCCCTTGTTCTTTGTCCTCGCCCTTTCTGAGAGTACTGTTCTGTAACTCTCAGCC 104483  
Qy 257 ----- 257  
Db 104482 CTTTACACGTGCAAGGAACCTTTAGTACTGTTGAGTCTTGAGACAACAATGCCTT 104423  
Qy 257 ----- 257  
Db 104422 ACTTCTCATAGCCGATTCGACAGGCCAGGTGTGTGCCAACCATGCTAGTTTAGACT 104363  
Qy 257 ----- 257  
Db 104362 CAGTTAGGAGGTGTTTGTGTCATGTATAATTCTCACTTCGAGTATACATGT 104303  
Qy 257 ----- 257  
Db 104302 GTTGAATCTGACTAGCTATATATTGCCAAAATTGTTACATTAATGCCATCTTAAT 104243  
Qy 258 ----- GlyThrAlaArgLeuL 263  
Db 104242 TATATGTCCTTATGATCTTAAATTCCTTTATGTTTTCAGGGAACAGCACATTTGA 104183  
Qy 263 euSerAsnSerPheAlaSerSerLeuTTPLeuMetProPheSerLybAnCybLySgLyA 283  
Db 104182 TGTTTAATGCAATTTGGTTCATCTTTATGCTCATGCTCCCTTCACGAATATTGCCAAGCG 104123  
Qy 283 sPAnThrSerTrpThrHisPhe ----- SerGlyGlyAla1AlaLybAspLybA 300  
Db 104122 ATATATATATACCTGAAGCATTTCTTGAAGGAAGGAGGTGGCCACCAAGACGACAA 104064  
Qy 300 rGVal1TyRHisCyAspGluGluTyRGlNserLybTySerGlyTyRProThrAsn1 320  
Db 104063 ----- TGTATGAATGAAATGAAATTAAGTCTGAATACACAGATGGCCCAAAACC 104015  
Qy 320 le11eAsn11eGlu11eProSerThr-SerVal1 ----- 330  
Db 104014 TTGTCAACATCGAAGTTCCTACAGTTCCAGGTTCTTTGTGAAATTTCTCCACACTTTA 103955  
Qy 331 ----- ThrGluThrAlaLeuVal1 ----- 336  
Db 103954 ATCTATGATTAAGATTTTGTCTCCACTGACGAGTCTACCTTTTCTTTCTTCAATTTTC 103895  
Qy 337 ----- AnMet 338  
Db 103894 TGTTTTCTCTGCGAGATACGGAACATACCATTCATTAATGATATCAACTGAGAACATA 103835  
Qy 339 Thr-SerMetGlyCySgLyLeuPProThrLeuLeuSerPheThrAlaArgGluLeuAla 357  
Db 103834 ACATCCAGCATGAGATGTGGAAGCAACTCTTTTGTCAATTTTCGCTATGGAGGTTTCA 103775  
Qy 358 AspGlyThrLeuPheLybAl11eGluAspTyRAspProAspSerLybArgMetLeuHis 377  
Db 103774 GATGTACTCTGTTTAAACCATTAAGATTTGGGACCTTCAGAGCATAGGGCTTATCCAT 103715  
Qy 378 GlNLeuLybLyLeu ----- 382  
Db 103714 CAGCTTGAGAAAGTG-AGTTTGCAATTTGTATGCCCAAGTGTCTACTTTGTAATTTGT 103656



AP008208\_049 4900001 5010000  
AP008208\_050 5000001 5110000  
AP008208\_051 5100001 5210000  
AP008208\_052 5200001 5310000  
AP008208\_053 5300001 5410000  
AP008208\_054 5400001 5510000  
AP008208\_055 5500001 5610000  
AP008208\_056 5600001 5710000  
AP008208\_057 5700001 5810000  
AP008208\_058 5800001 5910000  
AP008208\_059 5900001 6010000  
AP008208\_060 6000001 6110000  
AP008208\_061 6100001 6210000  
AP008208\_062 6200001 6310000  
AP008208\_063 6300001 6410000  
AP008208\_064 6400001 6510000  
AP008208\_065 6500001 6610000  
AP008208\_066 6600001 6710000  
AP008208\_067 6700001 6810000  
AP008208\_068 6800001 6910000  
AP008208\_069 6900001 7010000  
AP008208\_070 7000001 7110000  
AP008208\_071 7100001 7210000  
AP008208\_072 7200001 7310000  
AP008208\_073 7300001 7410000  
AP008208\_074 7400001 7510000  
AP008208\_075 7500001 7610000  
AP008208\_076 7600001 7710000  
AP008208\_077 7700001 7810000  
AP008208\_078 7800001 7910000  
AP008208\_079 7900001 8010000  
AP008208\_080 8000001 8110000  
AP008208\_081 8100001 8210000  
AP008208\_082 8200001 8310000  
AP008208\_083 8300001 8410000  
AP008208\_084 8400001 8510000  
AP008208\_085 8500001 8610000  
AP008208\_086 8600001 8710000  
AP008208\_087 8700001 8810000  
AP008208\_088 8800001 8910000  
AP008208\_089 8900001 9010000  
AP008208\_090 9000001 9110000  
AP008208\_091 9100001 9210000  
AP008208\_092 9200001 9310000  
AP008208\_093 9300001 9410000  
AP008208\_094 9400001 9510000  
AP008208\_095 9500001 9610000  
AP008208\_096 9600001 9710000  
AP008208\_097 9700001 9810000  
AP008208\_098 9800001 9910000  
AP008208\_099 9900001 10010000  
AP008208\_100 10000001 10110000  
AP008208\_101 10100001 10210000  
AP008208\_102 10200001 10310000  
AP008208\_103 10300001 10410000  
AP008208\_104 10400001 10510000  
AP008208\_105 10500001 10610000  
AP008208\_106 10600001 10710000  
AP008208\_107 10700001 10810000  
AP008208\_108 10800001 10910000  
AP008208\_109 10900001 11010000  
AP008208\_110 11000001 11110000  
AP008208\_111 11100001 11210000  
AP008208\_112 11200001 11310000  
AP008208\_113 11300001 11410000  
AP008208\_114 11400001 11510000  
AP008208\_115 11500001 11610000  
AP008208\_116 11600001 11710000  
AP008208\_117 11700001 11810000  
AP008208\_118 11800001 11910000  
AP008208\_119 11900001 12010000  
AP008208\_120 12000001 12110000  
AP008208\_121 12100001 12210000

AP008208\_122 12200001 12310000  
AP008208\_123 12300001 12410000  
AP008208\_124 12400001 12510000  
AP008208\_125 12500001 12610000  
AP008208\_126 12600001 12710000  
AP008208\_127 12700001 12810000  
AP008208\_128 12800001 12910000  
AP008208\_129 12900001 13010000  
AP008208\_130 13000001 13110000  
AP008208\_131 13100001 13210000  
AP008208\_132 13200001 13310000  
AP008208\_133 13300001 13410000  
AP008208\_134 13400001 13510000  
AP008208\_135 13500001 13610000  
AP008208\_136 13600001 13710000  
AP008208\_137 13700001 13810000  
AP008208\_138 13800001 13910000  
AP008208\_139 13900001 14010000  
AP008208\_140 14000001 14110000  
AP008208\_141 14100001 14210000  
AP008208\_142 14200001 14310000  
AP008208\_143 14300001 14410000  
AP008208\_144 14400001 14510000  
AP008208\_145 14500001 14610000  
AP008208\_146 14600001 14710000  
AP008208\_147 14700001 14810000  
AP008208\_148 14800001 14910000  
AP008208\_149 14900001 15010000  
AP008208\_150 15000001 15110000  
AP008208\_151 15100001 15210000  
AP008208\_152 15200001 15310000  
AP008208\_153 15300001 15410000  
AP008208\_154 15400001 15510000  
AP008208\_155 15500001 15610000  
AP008208\_156 15600001 15710000  
AP008208\_157 15700001 15810000  
AP008208\_158 15800001 15910000  
AP008208\_159 15900001 16010000  
AP008208\_160 16000001 16110000  
AP008208\_161 16100001 16210000  
AP008208\_162 16200001 16310000  
AP008208\_163 16300001 16410000  
AP008208\_164 16400001 16510000  
AP008208\_165 16500001 16610000  
AP008208\_166 16600001 16710000  
AP008208\_167 16700001 16810000  
AP008208\_168 16800001 16910000  
AP008208\_169 16900001 17010000  
AP008208\_170 17000001 17110000  
AP008208\_171 17100001 17210000  
AP008208\_172 17200001 17310000  
AP008208\_173 17300001 17410000  
AP008208\_174 17400001 17510000  
AP008208\_175 17500001 17610000  
AP008208\_176 17600001 17710000  
AP008208\_177 17700001 17810000  
AP008208\_178 17800001 17910000  
AP008208\_179 17900001 18010000  
AP008208\_180 18000001 18110000  
AP008208\_181 18100001 18210000  
AP008208\_182 18200001 18310000  
AP008208\_183 18300001 18410000  
AP008208\_184 18400001 18510000  
AP008208\_185 18500001 18610000  
AP008208\_186 18600001 18710000  
AP008208\_187 18700001 18810000  
AP008208\_188 18800001 18910000  
AP008208\_189 18900001 19010000  
AP008208\_190 19000001 19110000  
AP008208\_191 19100001 19210000  
AP008208\_192 19200001 19310000  
AP008208\_193 19300001 19410000  
AP008208\_194 19400001 19510000



Db	4902	ATATACATGCAATCTTTGCAAGTTGGTATGTAAACCAAGGAATTTAAATTCATCACATC	4843
Qy	232	-----GlyAlaProLeuLeuGlySerV	239
Db	4842	GTTATTTGTAGTGCATGCCATTTTTCCTGCTTTCAGTGCACCTCTTCTGGTTCTTA	4783
Qy	239	aIGuAlaIleIysSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu---	257
Db	4782	CTGAAGCAGTTTAAAGCTGCTCTTCTCGAGCAACATTTGGTCTTCAAGTCAGCGAGTCA	4723
Qy	257	-----	257
Db	4722	GCCTTTATGAATATATATATTTTCAATAATGTCTGCTGCGCTCAGAGGCTATATTTC	4663
Qy	257	-----	257
Db	4662	TGATATTTATACATTTTCTCCCAACATTTGTAAGTTGAATTAAGATGTGTGGTT	4603
Qy	257	-----	257
Db	4602	GTCGAATACTGAATAGTTTATTTTCCTTTTGATTAACATGATTTTGTGTATCTT	4543
Qy	257	-----	257
Db	4542	TTGCCCTTGTCTTTGTCCCTGCGCTTCTGAGATACGTTCGTGAACCTTCAGCC	4483
Qy	257	-----	257
Db	4482	CTTCTACACGTCAAGGAACCTTTAGTACTGTGGAGCTTTGAGACACAAATGCTT	4423
Qy	257	-----	257
Db	4422	ACTTCATAGCCGATTTGCAAGCGCAGGTGTGTGCACAAACATGCTAGTTAGAGT	4363
Qy	257	-----	257
Db	4362	CAGTTAGGAGGTGTTTGTGTCATGTAATAATTCTTACCTTGAGTATACATGT	4303
Qy	257	-----	257
Db	4302	GTTGAATCTGTACTTAAGTATATATGGCCAAATTTGTAACATTAATGCCATCTATAT	4243
Qy	258	-----GlyThrAlaArgLeuL	263
Db	4242	TATATGTCTCTTATGATCTAATAATTCCTTTTATTTGTTTTCAGGGAACAGCAGATGA	4183
Qy	263	euserFasSerPheAlaSerSerLeuTrpLeuMetProPheSerIysSerGlyTrpProThrAsnI	283
Db	4182	TGTTTAAATGCAATTTGGTTCATCTTATGCTCATGCCCTTCAGAAATATGTCCAAAGCTG	4123
Qy	283	sPAsnThSerItrPhrHisPhe-----SerGlyAlaIleAlaIleIysIlyAspIyVA	300
Db	4122	ATAATATATATCTGGAAGCATTTCTTTGAGGGAAGGGAGTTGCCACACACAGACCA-	4063
Qy	300	rgValIyHisCyAspGluGluGluIyTrpGlnSerIySerGlyTrpProThrAsnI	320
Db	4063	-----TGTATGAATATGAATATATAGCTGAATATCTCAGATATGCCACAAACC	4013
Qy	320	IeIleAsnIleGluIleProSerThr-SerVal-----	330
Db	4014	TTGTACAGCATCGAGGTTCTACAGATTCGAGGTTCTTTGTAATTTCTTCCAGCTTTA	3953
Qy	331	-----ThrGluThrAlaLeuVal-----	336
Db	3954	ATCTATGATTAAGATTTGCTTCACAGAGCTACCTTTTCTTTCAATTTTC	3893
Qy	337	-----AsnMet	338
Db	3894	TGTTTTCCCTGGCAGATACGGAAGCATCCCATCATTAATGATATCAACTGAGAAACA	3833
Qy	339	Thr---SerMetGluCySerGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuLeu	357

Db	3834	ACATCCAGCATGAGTGTGGAAAGCCAACTCTTTGGCATTTTCTGCTAGGAGGTTTCA	3775
Qy	358	AspGlyThrLeuPheIyValIleGluAspTryrAspProAspSerIyAsnMetLeuHis	377
Db	3774	GATGGTACTGTGTTTAAACCATTAAGATTGGGACCTGACAGACATAGGGCTTATCCAT	3715
Qy	378	GlnLeuIyValyLeu-----	382
Db	3714	CAGCTTGAGAGTGS-AGTTTGCATTTTGTATGCGCAGTGTCTGCTACTTTGTAAATTTGT	3656
Qy	382	-----	382
Db	3655	AATTCTCAGTTTCACATACATATTATTCTTCAAGTAACACCACTTACATGTTTTA	3596
Qy	382	-----	382
Db	3595	TATGCTTATGGAAGTAATAATGTGTGATTTTGTATTAATTGTATCTAATAACCCACACT	3536
Qy	383	-----TyrHisAspAspProValPheAsnProLeuThrPro	394
Db	3535	TTGATCTTTTGTGTGATCTCAGGTACTACAGGGTATCAGTTCATTAATCCCTCACACC	3476
Qy	395	TrpGluArgProPheIleIyAsnValPheCysIleIyGlyIaHisIeLeuIySethGlu	414
Db	3475	TGGAGAGACCCCAATTAAGATGTGTCTGCATATATGGCTTGATTCMAAGCTGAG	3416
Qy	415	Val-----	415
Db	3415	GT-ATGATGGGATTTTCTTGTATGTGTAAAGTTTCACTTTAGCTGCTTTGAAACAC	3357
Qy	416	-----GlyTyrTyrPheAlaPro	421
Db	3356	ATTACCATTCACGATTTATGCGCTTTTGTATGAGCAGTAGGCTATTAATTTTGCACCG	3297
Qy	422	SerGlyAspProIyProAspAsnTrpIleIleThrAspIleIleIyGluThrGluGly	441
Db	3286	AGTGAATAACATATCCAGATACGTATACGTATATACGTATATTATTATGAATTGAAAGG	3237
Qy	442	SerLeuValSerArg-----	446
Db	3236	TCTGTACATCAAGACGAGTATCATGATCTGTGTGAAAGTCTGTGTCCAAATCTT	3177
Qy	446	-----	446
Db	3176	TCAATATTCATTATCTGTCTTATGTATGTCTGTCTAGTGTAGTTCCTGGAAACG	3117
Qy	446	-----	446
Db	3116	TATTTTTTTTGTCTTAAATATGAAAAGATATTGTAGTGTGCGACATCTTTTTCATGGC	3057
Qy	447	-----SerGlyThr-ValIyAspGlyAsn-----	454
Db	3056	TAAATTGATTTGAATGAAGATCCTAGTAAAGTGCACGTGTTTCTAATAGTAAATCCACAG	2997
Qy	455	-----AlaGlyProIleThrGln	460
Db	2996	ATAATAGCTTTTGTGTGATTTGTCTTGTGAAAAATATCTGACTTTGAAACCTTGGCTTT	2937
Qy	460	YAspGlu-----	462
Db	2936	TGATGAGTTCAATTGTGTGTGGAACAACGTTTATCTGGCCACAATTTTATGTCTTTTGT	2877
Qy	462	-----	462
Db	2876	TTTAGAATAAATATGTAAACAACATTTATCTTCCAGATCAGTAAATCTGTACTGGA	2817
Qy	462	-----	462
Db	2816	AGCCCAACATTCACGCGAGACGGAACGATGTCTACTGTAGTATCTTTATAGTCA	2757
Qy	463	-----ThrAlaProIyHisSerLeuSeth	470
Db	2756	ATTCAATGCTTGTAGTTGACAAACAGACTTGTGTGCGAGGTATTCACAAATTCCTCTTC	2697

Qy 470 rTPCy9Ly9aSnTrpLeuGlyProLy9Val9aSnIleThrMet9AlaPro----- 486  
 |||||  
 Db 2696 ATGCTGACAGACACTGGCTTGGCCAAAGGAAACATACATAGGGCTCCGACGGATTTTA 2637  
 |||||  
 Qy 486 ----- 486  
 Db 2636 CCATGTGCTTGTATGTATGTATCTCCACAGAGAGATTAGGCTGTTTATTAATTTGTT 2577  
 |||||  
 Qy 486 ----- 486  
 Db 2576 TTGTGTGCTTATAGATTATGACTGTTCTTGTACATATCACAATGTAATGTC 2517  
 |||||  
 Qy 487 -----GlnProGluHisAspGlySerAspValHis 496  
 |||||  
 Db 2516 CTTTCCGCTTTGTATCTGACTGTTTTCGACGACGAGACATGATGATCTGATTTA 2457  
 |||||  
 Qy 496 sValGluLeuAsnValAspHisGlySerAspIleIleAlaAsnMetThrLysAl 516  
 |||||  
 Db 2456 AACCAAGATGATGTCACACACCATGTAGCGCAGGCGATCTCCCAACATGACGAGAC 2397  
 |||||  
 Qy 516 aProArgValIleTyrTleThrPheTyrGluAspSerGluSerIleProGlyLysArgTh 536  
 |||||  
 Db 2396 TCACATGTGAAGTACATACCTTGTGAAGATGCTGAAGCATTCGGGATGAGAAC 2337  
 |||||  
 Qy 536 rAlaValTrpGluLeuAspLysSerGly 545  
 |||||  
 Db 2336 AGCGTCTGGAGCTTGTATTAAGTGG 2309  
 |||||

RESULT 12  
 AP004120/c 118192 bp DNA linear PLN 02-JUN-2004  
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
 DEFINITION BAC clone:OJ1293\_E04.  
 AP004120  
 AP004120.3 GI:47847857

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Euphorbiales; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone:OJ1293\_E04  
 Published Only in Database (2001)  
 2 (bases 1 to 118192)  
 Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 Direct Submission  
 Submitted (29-AUG-2001) Takuji Sasaki, National Institute of  
 Agricultural Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:htp://rsgp.dna.affrc.go.jp/  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 On Jun 1, 2004 this sequence version replaced gi:34740243.  
 Genes were predicted from the integrated results of the following:  
 GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH  
 (http://www.softberry.com/), Genemark.hmm  
 (http://opal.biology.gatech.edu/genemark/), GlimmerM  
 (http://www.tigr.org/tdb/glimmer/glmr\_form.html), RiceHMM  
 (http://rsgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
 (http://bioinformatics.laastate.edu/cgi-bin/sp.cgi), sim4  
 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2  
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The  
 genomic sequence was searched against NCBI NonRedundant Protein  
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
 sequence database at RGP or DDBJ. Protein homologues of the coding  
 regions were searched against NCBI NonRedundant Protein database  
 with BLASTP. ESTs represent the identified cDNA sequences using  
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  
 Full-length cDNAs represent the identified cDNA sequences using  
 BLASTN with the corresponding DDBJ accession no.

## FEATURES

## source

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to JIGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from -21M3 to M3rev of the BAC clone. This sequence of OJ1293\_E04 clone has an overlap with OSUNBa0053h11 (DDBJ: AP005691) clone at 5' end and with p0643P09 (DDBJ: AP005111) at 3' end. The sequence was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at  
<http://rsgp.dna.affrc.go.jp/genomeseq.html>.

## Location/Qualifiers

1..118192  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39947"  
 /chr\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OJ1293\_E04"  
 /gene="OJ1293\_E04.1"  
 /join(<1863..1936,2024..>2468)  
 /gene="OJ1293\_E04.1"  
 /note="start and end point are not identified"  
 /join(1863..1936,2024..2468)  
 /gene="OJ1293\_E04.1"  
 /note="predicted by GeneMark.hmm etc."  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="BAD21651.1"  
 /db\_xref="GI:47847858"  
 /translation="MELRONCVQKLDGSKVPMKVPDFPFSEVVEFLKGYRVECGKI  
 PCMLGNSNVAIVFVVAEDSCGLASVVAIVGVGRQSTROVCLDSSNRDYYVE  
 GEIVPDGDSVILPQERADGWMELEGEFNPQGNNGVECFSLVPRKGRWLNSGCL  
 VIGIEIRPKIS"  
 /join(6482..6988,7071..7653,7919..8032,8112..8419,  
 8501..8573,8672..9312)  
 /gene="OJ1293\_E04.2"  
 /join(6482..6988,7071..7653,7919..8032,8112..8419,  
 8501..8573,8672..9312)  
 /gene="OJ1293\_E04.2"  
 /note="transposase-like"  
 /gene="OJ1293\_E04.3"  
 /join(11875..12242,12368..12486,13292..13972)  
 /gene="OJ1293\_E04.3"  
 /note="supported by full-length cDNA(s): AK060823"  
 /join(11953..12242,12368..12486,13292..13962)  
 /gene="OJ1293\_E04.3"  
 /note="supported by full-length cDNA(s): AK098981"  
 /join(11964..12242,12368..12486,13292..13787)  
 /gene="OJ1293\_E04.3"  
 /note="contains EST(s): AU092429(C50563)  
 contains full-length cDNA(s): AK060823,AK098981  
 phloem-specific lectin"  
 /codon\_start=1  
 /product="F-box family protein-like"  
 /protein\_id="BAD21652.1"  
 /db\_xref="GI:47847859"  
 /translation="MEEREGCLTEIARPELISAAISRPSPDCAHAASVAPFRA  
 ADDDAVASLPLNPLDLAGELSPAPSKKEIFLRISDGFYLSDBLKMWLDRERG  
 AKCYMISARLIVIMGTPTPYWMIPLTDSRFAGELELDVCMLELRGRTHSKLSFN  
 STYAAVWFPIADEFGIDAPFOEASVLSGRSTIKIVCSYDSDEBEPENYWPMS  
 IGPLRRARRRRRLVLDGVTVPQRTDEMMELEWGEFINSEGEDEVCFSIMETK  
 GNMKGLIYIGIEIRLKSG"



[illegible]

Db 65659 CATTTTCATACCTTGCACCTTAATCGGGTTCATGCTGACGCTCTTATCTGCTCAATTGC 65600  
Qy 89 TrpPheLeuCyseMetValIleuAspProGlyAsnGlnThrAspHisProGluCysLysSer 108  
Db 65599 TGGCTCAAAATGCATGCTGCTCGACCTTACACCAACGCGATCATCCAGAGTGCAGTCA 65540  
Qy 109 ArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGlyTyrIleThr----- 126  
Db 65539 AGGCTTCGACAGTGGTCTTCTCGCAATTACAGAGTTAGACCCTGGTTATATATACCTGGTAAA 65480  
Qy 126 ----- 126  
Db 65479 TATATGCTCTCGCTTAATCTCTTGGTATTGATTGATTTGAATTTTGAATACAAAT 65420  
Qy 126 ----- 126  
Db 65419 ATGCTGCACGATTCATCAGCAATGTGATAGATTGTGATTATGCTTCACTCCCATGCA 65360  
Qy 126 ----- 126  
Db 65359 TGGAAAAAAAACCATCGATGCGCAAGTGAGAGTGAATTAATTGACATCAAGCTTTTGT 65300  
Qy 126 ----- 126  
Db 65299 GAGTTAATAATGATATATGCCATGAGAAATAGCATTTAGCTTAGCTTACATTGCMAATTTC 65240  
Qy 126 ----- 126  
Db 65239 ATGTTCATCGGTAAAGCTGTAGAAATCCCATCTCGTTCCTCMAAGATTAATAAATGATT 65180  
Qy 127 -----GlyProLeu 129  
Db 65179 AGCCATTAATAAATCTTTCTTACTAATTTGGTTTATTTGGCCAAAGTTTTCAGGTCTCTCT 65120  
Qy 129 uSerThrValTrpLeuGlnTrpLeuLysTrpCysValGluPheGlyIleGluIleAsnAla 149  
Db 65119 TTCTTCAGTGTGAAAAAGAAATGGGTCAATAGTGTGTGAATTTGGTATGTGAAGCTAAATGC 65060  
Qy 149 aIleValAlaValProGlyAspTrpArgLeuSerProThrLysLeuGluIuArgAspLe 169  
Db 65059 CATTATGCGGTTCCATACATTCATTGAGGCTGCCCTTCATGCTTGAAGAGAGGAGATCT 65000  
Qy 169 uTyrPheHisLysLeuLys----- 175  
Db 64999 GTACTTCATTAATAATTAAAGTTTGTAAACATTCACCTCATATACATAGCAATCAATTAC 64940  
Qy 175 ----- 175  
Db 64939 TCACTGCCCATTTGGCTTCGATGATGATTGATTACTTAATTAATTTGGTGTAAAAACA 64880  
Qy 175 ----- 175  
Db 64879 CTTCTAAATTTGATCAACAGAAATATACCTACTTTGTTTACTCTCCATACCCTCAAAAA 64820  
Qy 175 ----- 175  
Db 64819 TCACGCAATACTAGTAGTACTAGTATCAGAAATACAAAGGTGTGGTGCCTTGGTGTG 64760  
Qy 175 ----- 175  
Db 64759 CATGCAAGCTGAAATGTATTAATTTTTTGGGGGGGAAATGTATTAAGCTTTTACTGCTT 64700  
Qy 175 ----- 175  
Db 64699 CTTTTCATTTTTTATGATCCTTTTCACCTTGAATTAATAAGTGTGTTTCTTCAAGTTTA 64640  
Qy 175 ----- 175  
Db 64639 TACAATACTGATATGACACGGCGATATGTCCGAAATTCATTTGTCANTCCGAAAGAT 64580  
Qy 176 -----LeuThrPheGluThrAlaLeuLysL 184  
Db 64579 AGTATTTTGTGCTCATATGCTCTTTTAACTGCAGGTTTACTTTTGAAACGTCAATTGAAC 64520

Qy 184 euArgGlyGlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArgTyrP 204  
Db 64519 TTTCAGAGAGGCGCTTCTTATAGTGTGCTCATTTCCATGGGTAAATATATGTTGGTACT 64460  
Qy 204 heLeuGlnTrpLeuArgLeuGluIleAlaProLysHisTyrLeuLysTrpLeuAspGlnH 224  
Db 64459 TTCTGGAATGTGTTGAACCTAGAAATGCTCCCAAGCATTTACATCCGATGGCTTGAAGAAC 64400  
Qy 224 IeIleHisAlaTyrPheAlaVal----- 231  
Db 64399 ATATACATCATACTTTCGACGTTGGATGTAAACCAAGAAATTAATTCAACCAATC 64340  
Qy 232 -----GlyAlaProLeuLeuGlySerV 239  
Db 64339 GTTATTTGTAGTGCATCCATTTTCTTCCTTCCATGCTTTCAGGTGACCTCTTCTGTTCTA 64280  
Qy 239 aIGluAlaIleLysSerThrLeuSerGlyValIlePheGlyLeuProValSerGlu---- 257  
Db 64279 CTGAAGCAGTTAAAGCTGCTCTTTCTGAGCAACATTTGGTCTTCAGTCAGCGAGTCA 64220  
Qy 257 ----- 257  
Db 64219 GCCTTATGAATATATATTTTCAATAATGTGCTCTGCCCTCAGAGGCTATATTTC 64160  
Qy 257 ----- 257  
Db 64159 TGATATTATACATTTTCTCCCAACATTTAGTGTGAATTAATGATATGTTGGTT 64100  
Qy 257 ----- 257  
Db 64099 GTCCAAATACTGAATAGCTTTTATTTCTTTTGGATTCTACAGTGTGTTGGTATCTT 64040  
Qy 257 ----- 257  
Db 64039 TTGCCCTGTCTTTTGCCCTGCCCTTCTGAGAGTCTGTTCTGTAACTCTGAGCC 63980  
Qy 257 ----- 257  
Db 63979 CTTCTACAGTCGACAGGAACCTTTTAGTACTGTGAGTCTTGGAGACAATAGCTT 63920  
Qy 257 ----- 257  
Db 63919 ACTTTCATAGCCGANTTCGACGGCCAGGTGTGTGCCACACCAATGCTAGTTTAGT 63860  
Qy 257 ----- 257  
Db 63859 CAGTTAGGAGGTTGTTTGGTGCATGTATAAATTCCTACCTTCGAGTGATACATGT 63800  
Qy 257 ----- 257  
Db 63799 GTTGAATCTGTACTAGCTATATATTGCCAAATTTGTACATTATATGCCATCTTAAT 63740  
Qy 258 -----GlyThrAlaArgLeuL 263  
Db 63739 TATATTGCTCTTATGCACTTAAATGCTTTTATTTGTTTTCAGGAGACGACGATTTGA 63680  
Qy 263 euSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCysLysGlyA 283  
Db 63679 TGTTTAAATGATTTGGTATCATCTTTATGCTCATGCTCTCTCAGAAATATGCAAAAGCTG 63620  
Qy 283 sPAsnThrSerTyrPheHisPhe-----SerGlyGlyAlaAlaLysLeuAspLysA 300  
Db 63619 ATATATATATCTGAGACATTTCTTTCAGAGGAAAGGAGGTTGCCACACAGACAGCAAG 63561  
Qy 300 rGlyAlaIleHisCysAspGluGluGlyTrpGlnSerLysTyrSerGlyTrpProThrAsnI 320  
Db 63560 -----TGTGATGAATGGAATATAGTCTGTAATCTCAGATGCGCCACAAACC 63512  
Qy 320 IeIleAsnIleGluIleProSerThr-SerVal----- 330  
Db 63511 TTGTACAGCATCAGAGTTCTTACAGTTCGAGGTTCTTGTGAAATTTCTTCCAGCTTTA 63452

```

QY      331 -----Thrg1uThrAlaLeuVal----- 336
Db      63451 ATCTATGATTAAGATTTTCTTCCTTCACACTGAGCAGCTACCTTTCTTTCTTTCAATTTC 63392
QY      337 -----AanMet 338
Db      63391 TGTTTTCTCTGGCAGATACGGAAAGCATCCCATCATGATGATACAACTGAGAAACATA 63332
QY      339 Thr---SerMetGluCyseGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAla 357
Db      63331 ACATCCAGATGAGAGTGGGAAAGCCAACTTTTGTCAATTTTCTGCTAGGAGGTTCGA 63272
QY      358 AaPpGlyThrLeuPheLeuAlaIleGluAaPyrAaPProAaPserLysArgMetLeuHis 377
Db      63271 GATGGTACTCTGTTTAAACATTAAGAGATTGGGACCTTCAGAGCATAGGCTTATCCAT 63212
QY      378 GlnLeuLysLeu----- 382
Db      63211 CAGCTTGAGAAAGG-AGTTGCAATTTTGATGCCAGAGTGTCTACTTTGTAATGTG 63153
QY      382 ----- 382
Db      63152 AATTCACAGTTTCACATACATATTATCTCAAGTAAACACAACTTACATGTTTTTA 63093
QY      382 ----- 382
Db      63092 TATGCTTATGAGTAATATGTGATTTTGATTAATTGTATCTTAATACCCCACTT 63033
QY      383 -----TyrHisAaPAspProValPheAaPProLeuThrPro 394
Db      63032 TTGATCTTTGTTGATCTCAGAGTACTACAGAGGTATCCAGATCTTAATCCCTCACACCC 62973
QY      395 TrpGluArgProProIleLysAenValPheCysIleTyrGlyAlaHisLeuLysThrGlu 414
Db      62972 TGGAGAGACCCCAATTAAGATGTCTGTGATATATGGGCTTGATTCAAAGACTGGAG 62913
QY      415 Val----- 415
Db      62912 GT-ATGATGGAGATTTTCTTGATTTGTTAAAGTTTCTACTTACCTGCTTGAAACAC 62854
QY      416 -----GlyTyrThrPheAlaPro 421
Db      62853 ATTACCATTCACCGATTATAGCCTTTTGTAAAGAGTAGGCTATTATTTTGACACGG 62794
QY      422 SerGlyLysProCysTyrProAaPAspAntPileIleThrAaPileIleLysGluThrGluGly 441
Db      62793 AGTGGAAACCATATCCAGATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTA 62734
QY      442 SerLeuValSerArg----- 446
Db      62733 TCTCTACTATCAAGCAGATATCATCTTGTGTGAAAGTCTCGTCCAAATCTT 62674
QY      446 ----- 446
Db      62673 TCAAATTATCCATTATCTGTCTTATGTATGTCTGCTTCTAGTATGATTCTCGAAGACAG 62614
QY      446 ----- 446
Db      62613 TATTTTTTTTGTCTTAATATGAAGAAAGATTTGTAGTTCGACACTCATTTTTCATGGC 62554
QY      447 -----SerGlyThrValValAaPLeuValAaPLeu 454
Db      62553 TAATTGATTTGAATGAAGATCTTAGTAAGTTGACATCGTGTTCATAGTGAATTAACACAG 62494
QY      455 -----AlaGlyProIleThrGly 460
Db      62493 ATAATAGCTTTTCGTGATTTGTTTTCTCGAAAAATATGACCTTGAACCTTTGCTTT 62434
QY      460 yAaPLeu----- 462
Db      62433 TGAATAGTTCAATTCGTGTTTGACAACTGTTTATCTGGCCACATTTTATGTCCTTTGT 62374
QY      462 ----- 462

```

```

Db      62373 TTTAGATTAATAAGTAAACAAATTATTTCTTCCAGATCAGTAATCTGTACTGGAA 62314
QY      462 ----- 462
Db      62313 AGCCCAACAATTCCAGCGGAGACGGAAAGGATAGTTCTACTGTAGTATCTTTAAGTCA 62254
QY      463 -----ThrValProTyrHisSerLeu 470
Db      62253 ATTCAATGCTTGAAGTTGAACAAACAGACTTTGTGGCAGAGTATCTTCAATTCCTCTC 62194
QY      470 rTrpCyseLysAntPLeuGlyProLysValAaPLeuThrMetAlaPro----- 486
Db      62193 ATGGCAAGAAACCTGCTGGGCCAAAGTGAACATACTAGGCTCCGACGATTTTA 62134
QY      486 ----- 486
Db      62133 CCATGTCCTAGATATGATTCCTCAGACAGAGAGATTGAGCCTGTTTATTAATTGTT 62074
QY      486 ----- 486
Db      62073 TTGTTGGCTATAGAAATTATGACTGTCTTCTTGACATATCAATGTAATGTC 62014
QY      487 -----GlnProGluHisAaPLeuSerAaPValHis 496
Db      62013 CTTTCCTGCTTTGTACATCTCTACTGTGTTTGGCAGGACAAACATGATGATGATTTACA 61954
QY      496 sValGluLeuAenValAaPHisLeuHisGlySerAaPLeuIleAlaAaPMetThrLysAl 516
Db      61953 MACCAAGATGAATGTGACACACATATGAGCCAGGCGCATCTCCAAACATGACAGAGAC 61894
QY      516 AProArgValLysTyrIleThrPheTyrGluAaPserGluSerIleProGlyLysArgTh 536
Db      61893 TCCACATGTAAGTACATTAACCTACTTTGAAGATCTGAGACATTCGGGATGAGAGAAC 61834
QY      536 rAlaValTrpGluLeuAaPLeuSerGly 545
Db      61833 AGCCGCTGGAGAGCTTGATTAAGGTGGG 61806

RESULT 13
AP005691/c 147123 bp DNA linear PIN 02-JUN-2004
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION BAC clone:OSUNBa0053111.
ACCESSION AP005691
VERSION AP005691.3 GI:47848524
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Echinoideae; Oryzaceae; Oryza.
REFERENCE
1 Sasaki,T., Matsunoto,T. and Katayose,Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSUNBa0053111
Published Only in Database (2002)
JOURNAL 2 (bases 1 to 147123)
AUTHORS Sasaki,T., Matsunoto,T. and Katayose,Y.
TITLE Direct SubMISSION
JOURNAL Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, url:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 1, 2004 this sequence version replaced gi:38564216.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://opal.biology.gatech.edu/Genemark/), Glimmer
(http://www.tigr.org/tdb/glimmer/glim_form.html), RicheHM
(http://rgp.dna.affrc.go.jp/RicheHM/), SplinePredictor
(http://bioinformatics.laestate.edu/cgi-bin/bp.cgi), sim4

```

(<http://globin.cse.psu.edu/html/docs/sima.html>), gap2 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (<http://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP or DDBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from -21M3 to M13rev of the BAC clone. This sequence of OSJNBa0053L11.1 clone has an overlap with P016F11 (DDBJ: AP005303) clone at 5' end and with OJ1293\_E04 (DDBJ: AP004120) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

## FEATURES

## source

```
1. 147123
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="genomic DNA"
    /cultiivar="Nipponbare"
    /db_xref="taxon:39947"
    /chromosome="2"
    /clone="OSJNBa0053L11.1"
    /join(210..282,431..528,603..662)
    /gene="OSJNBa0053L11.1"
    /join(210..282,431..528,603..662)
    /gene="OSJNBa0053L11.1"
    /note="hypothetical ORF
    predicted by GENSCAN
    this category is not included in IRGSP standard"
    2264..3226
    /gene="OSJNBa0053L11.2"
    /join(<2264..2293,2852..>3226)
    /gene="OSJNBa0053L11.2"
    /note="start and end point are not identified"
    /join(2264..2293,2852..3226)
    /gene="OSJNBa0053L11.2"
    /note="predicted by GENSCAN etc."
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="BAD22377.1"
    /db_xref="GI:47848525"
    /translation="MACHGDAVYRRRGTSARATESPPRRQASASARGSIASATSD
    SNAEGRLPLWLSASALSHSGVTSATIRNSYAVVAVVATSRGSPSLPPRATPPAAGYVA
    SPPRDTANSTRRGGTAAATRRPPLGSTGC"
    4436..6725
    /gene="OSJNBa0053L11.3"
    /join(4436..4894,4982..5164,5251..5600,6156..6617)
    /gene="OSJNBa0053L11.3"
    /note="supported by full-length cDNA(s): AK119257"
    /join(4438..4894,4982..5164,5251..5600,6156..6725)
    /gene="OSJNBa0053L11.3"
    /note="supported by full-length cDNA(s): AK063935"
    /join(4586..4894,4982..5164,5251..5600,6156..6333)
    /gene="OSJNBa0053L11.3"
    /note="contains EST(s): AU182147(C63648),AU058044(C63006)
    contains full-length cDNA(s): AK119257,AK063935"
    /codon_start=1
    /product="putative cinnamoyl CoA reductase"
    /protein_id="BAD22378.1"
    /db_xref="GI:47848526"
```

```
/translation="MSNSNMEANNDEKQEOVYCVTGAGGFTGSMVVRRLRGY
RVRGAPRPRKNAHLIDLEGAKERTLTCCADVLDPASLRAPAGCHGFPHASPPKSD
PNLVVAIAEGTRNNVKAADMGVRVRVTSVSGAVHMPNNSPDVLDSCSDSEPC
OREDIYCAKMAEKTATEASRRRLQALVAVVPCVTGPIIOPSVFSCHVVRVLTG
AAATYRNAVAAYADVDRARAHVLYEHNGARGRLCTGTVLRHAEILRMKLELPQY
PTVSGCEBGNQMVFPYKFSNQRRLDLGLEFPPLKSLHMAIECLQRKGLPVTVAQ
ORACL"
    complement(7826..8435)
    /gene="OSJNBa0053L11.4"
    /complement(join(<7826..8214,8309..>8435))
    /gene="OSJNBa0053L11.4"
    /note="start and end point are not identified"
    /complement(join(7826..8214,8309..8435))
    /gene="OSJNBa0053L11.4"
    /note="similar to Oryza sativa chromosome 1, P0497A05.2"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="BAD22379.1"
    /db_xref="GI:47848527"
    /translation="MAAGPSSAPPRPGAYATVYAGRGDSAAAGRGALHGRGA
    AVADVVERKLTATTAADCGSCRDAQPRHCGGERARDAFLHDASDPRVALHRSAP
    RDLVWRRLVALASDGGWTVGAVAVVAVAMVADGGGCGGGRYRCYLWRMLKA
    IIRLTVRHRRR"
    9814..11214
    /gene="OSJNBa0053L11.5"
    /join(<9814..10128,10328..10525,10605..10954,11058..>11214)
    /gene="OSJNBa0053L11.5"
    /note="start and end point are not identified"
    /join(9814..10128,10328..10525,10605..10954,11058..11214)
    /gene="OSJNBa0053L11.5"
    /codon_start=1
    /product="putative cinnamoyl CoA reductase"
    /protein_id="BAD22380.1"
    /db_xref="GI:47848528"
    /translation="MSISNDNNGDQKROOQOOLVCVTGAGGFTGSMVVRRLRGYR
    VYATVRDPDRKRAHLIALEGASRLSLERADVPAGLIAPAGCHGFVHACPLSN
    RQELVAVAVDGTGRVYMAAADMGVRVRVFTSSYCGAVHMPNNSPDVLDSCGWSPE
    FORKQDVSLHHVYCYAKTVAEMAAATEAKRGLEJLAVVPSMTGCMLORALNSTH
    VANYLYTGARKSYRNAVAAVYDVADVARALALYERDADAGRLCTGAVLHRLQLOML
    MDLPQYTIASKCDKDGKPMVKPYERSNRLDLGLEFPPLKSLYDAVMCMQRNHL
    PVLPL"
    complement(join(12759..12851,13439..13519))
    /gene="OSJNBa0053L11.6"
    /complement(join(12759..12851,13439..13519))
    /gene="OSJNBa0053L11.6"
    /note="hypothetical ORF
    predicted by GENSCAN
    this category is not included in IRGSP standard"
    15304..18533
    /gene="OSJNBa0053L11.7"
    /join(<15304..16325,17024..>18533)
    /gene="OSJNBa0053L11.7"
    /note="start and end point are not identified"
    /join(15304..16325,17024..18533)
    /gene="OSJNBa0053L11.7"
    /codon_start=1
    /product="HGW repeat containing protein-like"
    /protein_id="BAD22381.1"
    /db_xref="GI:47848529"
    /translation="MGVHVPADRLAFRAADWCFCLHGMPIIMPPLLGVDFTTGPPSP
    PTLVATYTGWSHHRGMASTSSPTGMPSPSPPTGVSAVYTGILLCHRCMASTSSPACIR
    RRLVLPPTKVAAYAAVAGRLRHRRPAFAADWCFRLHGMPIIMPPLLGVDFTTGAGS
    PSPSPGASATYTAGILCRRCMASTSSPPAYLRRRVVPPTRLVYAAVAGRLRHRRP
    ASVAWSGCLRHGMPIIMPPLLGVDFTTAGIPSPSPGASVYTLVGLLCRCRMASTSSIL
    AHLRCRLVPLRLTYVYAAISGLHFRHRRPSPSPPTGVFTVIDGLCSHIGFCHHQ
    GEYVKAQNGSVRTSPPARLDRLRVPPARLAMPPLLGVDFTTAGIPSPSPGASVYTL
    HGMPLRCRMASTSSPTGRVYRRRLVPPARLAMPPLLGVDFTTAGIPSPSPGASVYTL
    LHGMPIIMPPLLGVDFTTAGIPSPSPGASVYTLVAVVAVVATSRGSPSLPPRATPPA
    SACTAGILCRRCMASTSSPTGLSPSPPTGVSACTAGIATVAVRRLRHRLHRLACIRCL
    VPPPARLAMPPLLGVDFTTAGIPSPSPGASVYTLVPPARLVYAAVAGRLRHRRPACLR
    TGVSACTAGILCRRCMASTSSPTGRVYRRRLVPPARLAMPPLLGVDFTTAGIPSP
    CRVLPVPPARLAMPPLLGVDFTTAGIPSPSPGASVYTLVPPARLAMPPLLGVDFTTAGI
    PPTGVSACTAGIYAAVVGRLRHRRPACLRRLRVPPARLAMPPLLGVDFTTAGI
```

gene	FAAADWCRFHGMPIMLDLGVYVPTGKPRHPRLDVTASTAGLL"
	complement(18994. .19818)
	/gene="OSJUNBa0053l1.1.8"
	complement(join(<18994. .19134,19786. .>19818))
mRNA	/gene="OSJUNBa0053l1.1.8"
	/note="start and end point are not identified"
CDS	complement(join(18994. .19134,19786. .19818))
	/gene="OSJUNBa0053l1.1.8"
	/note="predicted by FGENSEH etc."
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="BAD22382.1"
	/db_xref="GI:47848530"
	/translation="MEVKKPVNSQICVEGVQVELINIGESLVAVKKHGICRQVEVH
	KVVVGEVVKLAK"
gene	complement(join(21959. .22112,22149. .22168))
	/gene="OSJUNBa0053l1.1.9"
	complement(join(21959. .22112,22149. .22168))
misc_feature	/gene="OSJUNBa0053l1.1.9"
	/note="hypothetical ORF
	predicted by GlimmerM
	this category is not included in INRSP standard"

Alignment Scores:	
Pred. No.:	1,05e-115
Score:	1483.50
Percent Similarity:	29.50%
Best Local Similarity:	26.17%
Query Match:	50.03%
DB:	15
	Gaps:
	17
US-09-651-651-5 (1-546) x AP005691 (1-147123)	
	Length: 147123
	Matches: 401
	Conservative: 51
	Mismatches: 81
	Indels: 1001

[illegible]

Db 126797 CTTTCATTTTTTATGACATCCTTTTCATGAAATTAAGTTTCTTCAAGTTTA 126738  
 Qy 175 ----- 175  
 Db 126737 TACAACTACTGTAATATGACACGGCCGATATGTCGAAATTCATTGTCATGGCGAAAGAT 126678  
 Qy 176 -----LeuThrPheGluThrAlaLeuLysL 184  
 Db 126677 AGATTTTGGTGTCTCATAGTCCTTTTAACGACGTTAACTTTTGAACCTGACATGAAAC 126618  
 Qy 184 euArgGlyGlyProSerLLeuAlaPheAlaHisSerMetGlyAsnAsnValPheArgTyrP 204  
 Db 126617 TTCAGAGAGGGCCTCTCTTATGTTGTGCTCATTCATGGGTATATATGTTGTTGGTACT 126558  
 Qy 204 heLeuGluThrPheuArgLeuGluLeuAlaProLysHisTyrLeuLysLThrPheuAspGlnH 224  
 Db 126557 TTCCTGAATGTGGAACATAAGAAATGCTCCCAAGCATTAATCCATGGCTTGACGAAAC 126498  
 Qy 224 IeIleHisAlaTyrPheAlaVal----- 231  
 Db 126497 ATATACATGCATCTTTCGAGTTGTATGTAACAGAAAGAAATTAATTCATCACCATTC 126438  
 Qy 232 -----GlyAlaProLeuLeuGlySerY 239  
 Db 126437 GTTATTTGTTAGTGTATGCCATTTTTCCTGCTTCAGGTGACCTCTTCTGTTCTTA 126378  
 Qy 239 aIGluAlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGly---- 257  
 Db 126377 CTGAACGATTAAGAGCTGCTCTTCTGAGCAACATTTGTCCTTCAGTCAGCAGGTCA 126318  
 Qy 257 ----- 257  
 Db 126317 GCCTTATGATATATATATTTTCAATAATGTGTCCTGCTCAGAGAGCTATATATTC 126258  
 Qy 257 ----- 257  
 Db 126257 TGATATTATATACATTTTTCCTCAACATGTGTAGTGAATTAAGTATGTGTGTT 126198  
 Qy 257 ----- 257  
 Db 126197 GTCCAACTATGAAATAGTTTATTTCTTTTGGATTCACTAGTTTGTGTGATTTCTT 126138  
 Qy 257 ----- 257  
 Db 126137 TTGCCCTTGTCTTTTGTCCCTGCTTCTGAGAGTACTGTCTGTAACTCTCAGACC 126078  
 Qy 257 ----- 257  
 Db 126077 CTCTACACGTCGCAAGGAACCTTTTAGTACTGTGAGTCTTGAGACAACAATGCTT 126018  
 Qy 257 ----- 257  
 Db 126017 ACTTTCATAGCCGATTGCAAGCCAGGTGTGTGCCACAACCAAGCTAGTTTGAAGT 125958  
 Qy 257 ----- 257  
 Db 125957 CAGTTAGGAGGTTGTTTGGTGATGATTAATTCCTACCTTCGAGTGATACACATGT 125898  
 Qy 257 ----- 257  
 Db 125897 GTTGAATCTGTACTTACTATATATTCGCAAAATTTGTAATTAATGCACTTAAAT 125838  
 Qy 258 -----GlyThrAlaArgLeuL 263  
 Db 125837 TATATGTCTCTTATGATCTAAATATGCTTTTATGTTTTCAGAGGAACAGCAGATTTGA 125778  
 Qy 263 euSerAsnSerPheAlaSerSerLeuThrPheuMetProPheSerLysAsnGlyLysGlyA 283  
 Db 125777 TGTATTATGATTTGGTTCATCTTTATGGCTCATGCCCTTCGAAATATGCAAAAGCTG 125718  
 Qy 283 sPheThrSerTyrThrHisPhe-----SerGlyGlyAlaAlaLysLysAspLysA 300

Db 125717 ATATATATATCTGAGAGATTTCTTTGAGGGAAGGAGGTTGCCACCAAGACAGCA- 125659  
 Qy 300 rGValTyrHisCyAspGluGluGlyTyrGlnSerLysTyrSerGlyTyrProThrAsnI 320  
 Db 125658 -----TGTGATGAATATGAAATATAGTCTGAATATCTCAGATAGCCACAAACC 125610  
 Qy 320 IeIleAsnIleGluIleProSerThr-SerVal----- 330  
 Db 125609 TTGTGACATTCGAGGTCTCTTACAGTTGCAAGTTCCTTTGTAATTTCTTCCACCTTTA 125550  
 Qy 331 -----ThrGluThrAlaLeuVal----- 336  
 Db 125549 ATCTATGATTAAGATTTTGTCTTCCACTGACGACGTCCTACCTTTCTTTTCTTCAATTTTC 125490  
 Qy 337 -----AsnMet 338  
 Db 125489 TGTTTTCTCTGCGAGATACGGAACATPACCATCATTAATGATATCACTGAGAACATA 125430  
 Qy 339 Thr---SerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAla 357  
 Db 125429 ACATCCAGCATGAGGTGTGGAAGCCACCTTTTGTGCTATTTCTGCTAGGAGGTTTCA 125370  
 Qy 358 AspGlyThrLeuPheLysAlaIleGluAspTyrAspProAspSerLysArgMetLeuHis 377  
 Db 125369 GATGTACTCTGTTTAAACATTAAGATTTGGACCTCAGACGATAGGCTTATCCAT 125310  
 Qy 378 GlnLeuLysLysLeu----- 382  
 Db 125309 CAGCTTGAAAGATG-AGTTGCAATTTGTATGCCGACAGTCTGCTACTGTGAATGT 125251  
 Qy 382 ----- 382  
 Db 125250 AATTCAGTTTCACATCATATTAATCTTCAAGTAACCAACCAACTTAATGTTTTA 125191  
 Qy 382 ----- 382  
 Db 125190 TATGCTTATGAGAAATATATGTGATTTGATTAAATTTGATCTAATACCCCACTT 125131  
 Qy 383 -----TyrHisAspAspProValPheAsnProLeuThrPro 394  
 Db 125130 TTGATCTTTGTGATCTCAGTACTACACAGGAGATCCAGTTCTTATATCCCTCAGACCC 125071  
 Qy 395 TrpGluArgProProIleLysAsnValPheCysIleTyrGlyAlaHisIleuLysThrGlu 414  
 Db 125070 TGGAGAGAGCCCCCAATAAAGATGTCTGCATATATGAGCTTGATTCAAAGACTGAG 125011  
 Qy 415 Val----- 415  
 Db 125010 GT-ATGATTTGGATTTTCTTTGATTGTTAAAGTTTCTACTTTAGCCTGCTTGAACAC 124952  
 Qy 416 -----GlyTyrTyrPheAlaPro 421  
 Db 124951 ATTACATTCACCGATTATATGCGCTTTTGTAAAGCAGTGAAGCTATATTTTGACCG 124892  
 Qy 422 SerGlyLysProTyrProAspAsnTrpIleIleThrAspIleIleTyrGluThrGluGly 441  
 Db 124891 AGTGAAAAACCATATCCAGATTAATCTGATTAATTAACGATTAATTAATTAATTTGAAGG 124832  
 Qy 442 SerLeuValSerArg----- 446  
 Db 124831 TCTCTACTATCAAGGACGATATCATGATCTGTGTGGAAAGCTTCGTGCCAATTTCTT 124772  
 Qy 446 ----- 446  
 Db 124771 TCAATTAATCAATTATCTGCTTATGTATGTCTGCTCTAGTGTAGTTTCTGGAACAG 124712  
 Qy 446 ----- 446  
 Db 124711 TATTTTCTTCTTAAATATGAAAAAGATTTGATGTCGCACTCATTTTTCATGGC 124652  
 Qy 447 -----SerGlyThr-ValIleAspGlyAsn----- 454  
 Db 124651 TAATTTGATTGAATGAAGATCCTAGTAAGTTGACGCTTTTCTAATAGTAAATTAACACAG 124592

QY	455	-----	AlaGlyProIleThrG1	460
Db	124551	ATAAATAGCTTTTCGTCGATTTGTTTCTCGAAAAAATACTGACCTTGGAACTTTGGCTTT	124532	
QY	460	YAPGlu-----	462	
Db	124531	TGATGAGTTCAATTCGTGTTTGGAACAATGTTTATCTGGCACAATTTTATGTCCTTTGGH	124472	
QY	462	-----	462	
Db	124471	TTTAGAATAAATAGTAAACAATTTATTCCTTCAGATCAGGTAAATCTGTTACTGGAA	124412	
QY	462	-----	462	
Db	124411	AGCCCAACAATTCACGCGAGACGGAACGGTATGTTCTACTGATGATCTTTATAGTCA	124352	
QY	463	-----ThrValProIleHisSerLeuSe	470	
Db	124351	ATTCAATGTCCTTGAGTTGAACAACAAGACTTTGTGGCAGGTATCCTACAAATCTCTCTC	124292	
QY	470	TTTCybelysaerTPLeuGlyProIlySValanIleThrMetAlaPro-----	486	
Db	124291	ATGTGTCAAGAACTGCTGTGGCCAAAGTGAACATTAAGGCTCCGCGAGTATTTTA	124232	
QY	486	-----	486	
Db	124231	CCATGTCCTAGTGAATATGATCTCTACGACAGAGAGTTTGAGGCTGTTTATTAATTTGTT	124172	
QY	486	-----	486	
Db	124171	TTGTTGTGCTAATAGAAATTAATGACTGTTTCTTGACATATCAATATGTAATGTGC	124112	
QY	487	-----GlnProGlnHisAspGlySerAspValIle	496	
Db	124111	CTTTCTGCTGTTGTATCATCTCTACTGTTTGTTCAGGCAGAAACATGATGATCTAATTTACA	124052	
QY	496	evAlGluLeuAenValIAspHisGlyIleSerAspIleIleAlaAsnMetThrIlyVal	516	
Db	124051	AACAAGATGATGTGACACCAACATGATGAGCCAGCGCATCTCCCAACATGACGAGGAC	123992	
QY	516	aProArgValIlyeTyrlleThrPheTyrgIuaAspSerGluSerIleProGlyIlysaIgh	536	
Db	123991	TCCACATGTGAAGTACATAACCTTATGGAAGATGCTGAGAGCATTCGCGGATGAGGAAAC	123932	
QY	536	raIaValITpGluLeuAspIlySerGly	545	
Db	123931	AGCCGTCGAGACTGATTAAGCTGGG	123904	
RESULT 14				
AP004882/c		158971 bp	DNA	linear
LOCUS		Oryza sativa (japonica cultivar-group)	chromosome 2	clone P0511E12,
DEFINITION		*** SEQUENCING IN PROGRESS ***.		
ACCESSION		AP004882		
VERSION		AP004882.1		
KEYWORDS		GI:19698302		
SOURCE		HTG: HTGS PHASE2.		
ORGANISM		Oryza sativa (japonica cultivar-group)		
		Oryza sativa (japonica cultivar-group)		
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
		Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE		1		
AUTHORS		Sasaki, T., Matsumoto, T. and Yamamoto, K.		
TITLE		Oryza sativa nipponbare (GH3) genomic DNA, chromosome 2, PAC		
JOURNAL		clone: P0511E12		
REFERENCE		2		
AUTHORS		Published Only in Databae (2002)		
TITLE		2 (bases 1 to 158971)		
JOURNAL		Sasaki, T., Matsumoto, T. and Yamamoto, K.		
		Submitted (20-MAR-2002) Takuji Sasaki, National Institute of		
		Agricobiological Sciences, Rice Genome Research Program; Kamondai		
		2-1-2, Tsukuba, Ibaraki 305-8602, Japan		

TE-mail:tsasak@nias.affrc.go.jp, URL:htp://xsp.dna.affrc.go.jp/  
TEL:81-298-38-7441, Fax:81-298-38-7469)

COMMENT:

NOTE: It currently consists of 1 contig. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it  
available and the accession number will be preserved.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

Location/Qualifiers

ORIGIN

Alignment Scores:	
Pred. No.:	1,15e-115
Score:	1483.50
Percent Similarity:	29.50%
Best Local Similarity:	26.17%
Query Match:	50.03%
OB:	14
Length:	158971
Matches:	401
Conservative:	51
Mismatches:	81
Indels:	1001
Gaps:	17

US-09-651-651-5 (1-546) x AP004882 (1-158971)

[illegible]



Db 59464 GTAAGCGTCATACGGCTTAATTGGGTGAGTTGTTACAGAACATTCATCTGTCAAT 59405  
 QY 81 ----- 81  
 Db 59404 GAAACGGCTGTAAAGATTATTGTTACATGAGGGGTCATATTAGTATTATCTTT 59345  
 QY 82 ----- LeuSerAlaValAenCys 88  
 Db 59344 CATTTCATACCTTGCACCTTAATCGGGTTCATGTGACGCTCTTATCTGTCAATTGC 59285  
 QY 89 TrpPheLeuCyMetValLeuAspProTyrAspGlnThrAspHisProGluCysIysSer 108  
 Db 59284 TGGCTCAATAGCATGCTGCTCGAACCTTTACACACGAGATCATCCAGAGTGCAGTCA 59225  
 QY 109 ArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGlyTyrIleThr----- 126  
 Db 59224 AGGCTGACAGTGGTCTTCTGTCAATTACAGAGTTAGACCTTGTTATTATACCTGGTAA 59165  
 QY 126 ----- 126  
 Db 59164 TTATATGCTCGCTTAATTCCTTGTTGATTGATTGATTGATTGATTGATTGATTGATT 59105  
 QY 126 ----- 126  
 Db 59104 ATGCTGCACGATTCATCAGCAATGTGATAGATTGTTATGCTTCACCTCCCATGCA 59045  
 QY 126 ----- 126  
 Db 59044 TGGAAAAAAACCATCGATGCGCAAGTGAAGTGAATAATTGACATCAAGCTTTGT 58985  
 QY 126 ----- 126  
 Db 58984 GAGTTAATATGATATTGCCATGAGATAATAGCATTAGCTAGTACATTGCAATTC 58925  
 QY 126 ----- 126  
 Db 58924 ATGTTTCATCGTAGTGAATTCCTCCCATCTGTTCTCAAGAAATAAAATGATT 58865  
 QY 127 ----- GlyProLe 129  
 Db 58864 AGCCATAAAAAATCTTTCTTACTAATTGGTTTATTTGGCCAAAGTTTTCAGGTCCTCT 58805  
 QY 129 uSerThrValTyrIysGluTyrPheLeuLysTyrCysValGluPheGluIleGluAlaAsnAl 149  
 Db 58804 TTCTTCAGTGTGAAAGATGGGTCAATGTGTGTGAATTGTGTAATTGAGCTAATGC 58745  
 QY 149 alleValAlaValIProTyrAspTTPArgLeuSerProThrLysLeuGluGluAlaGAspLe 169  
 Db 58744 CATTATGCGGGTTCATACGATGTGAGGCTGCCCTTCATGCTTGAGGAGAGGATCT 58685  
 QY 169 uTyrPheHisIleLysLeuLys----- 175  
 Db 58684 GTACTTCCATAAATTAAAGTTTGTAAACCTTCACTCATATACATAGCAATCAATTAC 58625  
 QY 175 ----- 175  
 Db 58624 TCACTGCCCATGTGGCTTGCATAGATTGTGATTATTACTAATAAATTGGTAAAAACA 58565  
 QY 175 ----- 175  
 Db 58564 CTTCATAATTGTATCAACAGATAATACTACTTTGTTTACCTTCATACCTCAAAAAA 58505  
 QY 175 ----- 175  
 Db 58504 TCACGCATACTAAGTACTAGTATGCAAAATAAAGGTTGGTTGGCTTTGGGTG 58445  
 QY 175 ----- 175  
 Db 58444 CATGAAGCTGAAATTGATTAATTTTTTTTGGGGGGAATTGATTAAGCTTTTACTTGCTT 58385  
 QY 175 ----- 175

Db 58384 CTTTTCATTTTTTTATGACATCTTTTCACTTGAAATTAAGTGTGTTTCTTCAAGTTTA 58325  
 QY 175 ----- 175  
 Db 58324 TACAATACTTGAAATATGACACGGCGATATGCGAAATTCATGTGTCATGCGGAAAGAT 58265  
 QY 176 ----- LeuThrPheGluTyrAlaIleLysVal 184  
 Db 58264 AGTATTGTTGGTCTCATAGTCTCTTTTAACTGACAGGTTAACTTTTGAAACTGCATTGAAAC 58205  
 QY 184 euArgGluYsIProSerIleValIlePheAlaHisSerMetGlyAsnAsnValIleArgTyrP 204  
 Db 58204 TTCGAGAGGGCTCTTCTTATAGTGTGTCATTTCCATGGGTATATATAGTGTTCGTTACT 58145  
 QY 204 heLeuGluTyrPheLeuArgLeuGluIleAlaProLysHisTyrLeuLysTyrPheLeuAspGlnH 224  
 Db 58144 TTCGGAATGTTGGAACCTAGAAATGCGTCCCAAGCATTAATCCAGTGGCTTGACGAAC 58085  
 QY 224 IleIleHisAlaTyrPheAlaVal----- 231  
 Db 58084 ATATATCATGCATATCTTGTGACGTTGTATGTAAACAGAAATTTAATTCATCACCATTC 58025  
 QY 232 ----- GlyAlaProLeuLeuGlySerY 239  
 Db 58024 GTTATTGTTAGTGCATGCCATTTTCTTGCTTTCAGGTGACCTCTTCTTGTTCTTA 57965  
 QY 239 aIleGluAlaIleLysSerThrLeuSerGlyValIleThrPheGlyLeuProValSerGlu----- 257  
 Db 57964 CTGAACAGTTAAAGCTGCTCTTCTGAGCAACATTTGGTCTTCAGTCAGCAGAGTCA 57905  
 QY 257 ----- 257  
 Db 57904 GCCTTTGAATTAATTAATTTTCAATTAATGTCTCTGCGCTCAGAGGCTAATTTTC 57845  
 QY 257 ----- 257  
 Db 57844 TGATATTATACATTTTCTCCCAATGTTAGTGTGAATTAAGTATGTTGTTGTT 57785  
 QY 257 ----- 257  
 Db 57784 GTCCAAATACGAATAGTTTATTTCTTTTGGATTCACTAGGTTTGTGTGATTTCTT 57725  
 QY 257 ----- 257  
 Db 57724 TTGCCCTGTGTTCTTTGTCCCTGCTCTTCTGAGAGTACTGTTCTGTAACTCTCAACC 57665  
 QY 257 ----- 257  
 Db 57664 CTTTACACGTGCAAGGAACCTTTTATGACTGTGGAAGTCTTGAGACAACAATGCTT 57605  
 QY 257 ----- 257  
 Db 57604 ACTTTCATAGCCGATTGCGACGGCAGGTGTGTGCCAACAAACATAGTATTAGAGT 57545  
 QY 257 ----- 257  
 Db 57544 CAGTTAGGAGGTGTTTGGTGCATGTATTAATTTCTTCACTTGCAGTATACACATGT 57485  
 QY 257 ----- 257  
 Db 57484 GTTGAATCTGTACTTACCTATATATGCCAAATTTGTACATTATATGCCATCTTAAAT 57425  
 QY 258 ----- GlyThrAlaArgLeuL 263  
 Db 57424 TATATGTCTTATGCACTAAATTTGCTTTATGTTTTCAGGGAACAGACAGATTGA 57365  
 QY 263 euSerAsnSerPheAlaSerSerLeuThrLeuMetProPheSerLysAsnCysAlaArgIle 283  
 Db 57364 TGTTTATGCAATTTGGTTCATTTTATAGCTCAAGCTCTTCTGAAATATATGCAAGCTG 57305  
 QY 283 sPheAsnThrSerTyrThrHisPhe-----SerGlyGlyAlaAlaLysLysAspLysA 300  
 Db 57304 ATTAATATATCTGGAACATTTCTTGAAGGAAAGGAGTGTGCACACAGACAGCAA- 57246

QY 300 rgValYrhiScYsApGluGluYrGlnSerLySerGlyYrProThrAsnI 320  
 Db 57245 -----TGTGATAAAAGGAATATAGCTGAAATACAGATGGCCACAAACC 57197  
 QY 320 leIleAsnIleGluIleProSerThr-SerVal----- 330  
 Db 57196 TTGTGACATCGAGGTTCTTCACTGACGAGTTCCTTTGTAATTTTCCACGCTTAA 57137  
 QY 331 -----ThrGluThrAlaLeuVal----- 336  
 Db 57136 ATCTATGATTAGATTTTGTCTTCCACTGACGAGTCTACCTTTTCTTTCTTCAATTTTC 57077  
 QY 337 -----AsnMet 338  
 Db 57076 TGTTTTCTGTGACAGATACGGAAGCATACCATCATTTATGATACAACTGAGAACATA 57017  
 QY 339 Thr-----SerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgIleuAla 357  
 Db 57016 ACATCCAGCATGAGTGTGGAAGCACTTTTGTCTATTTTCTGTCTGAGGAGGTTTCA 56957  
 QY 358 AspGlyThrLeuPheLeuValIleGluAspTyrAspProAspSerLeuValArgMetLeuHis 377  
 Db 56956 GATGGACTCTGTGTTAAACATTAAGATTGGACCTCAGAGCATAGGCTTATCCAT 56897  
 QY 378 GlnLeuValLeuLeu----- 382  
 Db 56896 CAGCTTGAGAGATG-AGTTTGCAATTTTGTATGCCAGCTGTCTACTTTGTAATTTGT 56838  
 QY 382 ----- 382  
 Db 56837 AATTGCACTTCACATACATATTTATCTTCAAAAGTAAACCAACCACTTACATGTTTAA 56778  
 QY 382 ----- 382  
 Db 56777 TATGCTTATGGAATATATGATGATTTTGATTAATTGATCTAATACCCACATTT 56718  
 QY 383 -----TyrHisAspAspProValPheAsnProLeuThrPro 394  
 Db 56717 TTGATCTTTTGATGCTCAGGATCTACGAGGATGATCCAGTTCTTAAATCCCTCACAACCC 56658  
 QY 395 TrpGluArgProProIleLeuValAsnValPheCysIleTyrGlyAlaHisLeuValThrGlu 414  
 Db 56657 TGGGAGAGACCCCATTAAGATGTCTGTGATATATGGCTTATTCAAAGACTGAG 56598  
 QY 415 Val----- 415  
 Db 56597 GT-ATGATTGGAGATTTTCTTGTATTGTTAAAGTTTCTACTTTAGCCTGCTTGAACAC 56539  
 QY 416 -----GlyTyrTyrPheAlaPro 421  
 Db 56538 ATTACCATTCACCGATTTATGCGCTTTTGTATAGCAGGTAGGCTAATATTTTGGACCG 56479  
 QY 422 SerGlyLeuProTyrProAspAsnThrIleIleThrAspIleIleTyrGluThrGluGly 441  
 Db 56478 AGTGGAAAACCATATCCAGATTAAGTGAATTAAGTGAATTTTATGAAATTTGAAGG 56419  
 QY 442 SerLeuValSerArg----- 446  
 Db 56418 TCTCTACTATCAAGGACGATATCATCTTGTGTGAAAGTCTCGTCCAAATCTT 56359  
 QY 446 ----- 446  
 Db 56358 TCMAATTAATCCATTATCTGCTTTATGATGTCGTCTTCTAGTGAATTTCTCGAAGACG 56299  
 QY 446 ----- 446  
 Db 56298 TATTTTCTTGTCTTAATATGAAGATTTTGTAAGTTCGACGACATCTTTTCATGGC 56239  
 QY 447 -----SerGlyThr-ValValAspGlyAsn----- 454  
 Db 56238 TAATTTGATTGAATGAAGATCTTAGTAAGTTTGACATGGTTTCTAATGTTAATTAACACAG 56179

QY 455 -----AlaGlyProIleThrGln 460  
 Db 56178 ATATAGCTTTTTCGTGATTTGTTTCTCGAAAAATAGTACCTTTGAACCTTTGGCTTT 56119  
 QY 460 YAspGlu----- 462  
 Db 56118 TGATGAGTTCAATTCGTGTTTGACAACAGTGTATCTGGCCACATTTTATGTCCTTTGT 56059  
 QY 462 ----- 462  
 Db 56058 TTTAGATAATAATAGTAACACAAATTTATCTTCAGATCAGTAATTTCTGTACTGAA 55999  
 QY 462 ----- 462  
 Db 55998 AGCCCAACAATTCACGCGGAGACGGAACGGATGTCTTACTGTATCTTATAGTCA 55939  
 QY 463 -----ThrValProTyrHisSerLeu 470  
 Db 55938 ATTCAATGCTTGAGTTGACAAACAAAGACTTTGAGCAGGATATCTCAATTCCTCTC 55879  
 QY 470 rTrpCysAluValAsnThrLeuGlyProValAsnIleThrMetAlaPro----- 486  
 Db 55878 ATGGTGCAAGAACTGGCTTGGCCAAAAGTGAACATACTAGGCTCCGACAGTATTTTA 55819  
 QY 486 ----- 486  
 Db 55818 CCATGCTGTAGTATGATATCTTCAGACAGAGATTGAGGCTTTATTAATTTGT 55759  
 QY 486 ----- 486  
 Db 55758 TTGTTGGCTAATGAAATTAATGACTTTCTGTAGCATATCACAATGTAATGTGC 55699  
 QY 487 -----GlnProGlnHisAspGlySerAspValHis 496  
 Db 55698 CTTTCTGCTTTGATCATCTTACTGTTTTCAGACCAACATGATGATCTGATTTCA 55639  
 QY 496 ValGluLeuValAsnValAspHisGluHisGlySerAspIleIleAlaAsnMetThrLeuAl 516  
 Db 55638 AACAAAGATGAATGTGACACACATGATAGCCAGGCACTATCTCCAAACATGACAGGAC 55579  
 QY 516 AsProArgValIleTyrIleThrPheTyrGluAspSerGlnSerIleProGlyValArgTh 536  
 Db 55578 TCCACATGAGAAATGATCAATCACTTGAAGATCTGAGAGCATTCGCGATGAGAAC 55519  
 QY 536 rAlaValAlrProGluLeuAspGlySerGly 545  
 Db 55518 AGCCGCTGGAGCTTGATTAAGGTGG 55491  
 RESULT 15  
 AP006356/c 93398 bp DNA linear PLN 22-JUL-2003  
 LOCUS AP006356  
 DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 6,  
 clone:J1706108, TM0082b, complete sequence.  
 ACCESSION AP006356  
 VERSION AP006356.1 GI:31580987  
 KEYWORDS HTG.  
 SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)  
 ORGANISM Lotus corniculatus var. japonicus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;  
 Lotus.  
 REFERENCE 1  
 ASAMIZU, E., KATO, T., SATO, S., NAKAMURA, Y., KANEKO, T. and TABATA, S.  
 Structural Analysis of a Lotus japonicus Genome. IV. Sequence  
 Features and Mapping of a seveny-three TAC clones which cover the  
 7.5 Mb Regions of the Genome  
 DNA Res. (2003) In press  
 JOURNAL 2 (bases 1 to 93398)  
 AUTHORS Sato, S.  
 TITLE Direct Submission  
 REFERENCE Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,  
 Department of Plant Gene Research, 2-6-7 Kazusa-kamatori, Kisarazu,

Chiba 292-0818, Japan (E-mail: sato@kazusa.or.jp)  
 URL: http://www.kazusa.or.jp/ Tel: 81-438-52-3935 (ex. 2337),  
 Fax: 81-438-52-3934)

## FEATURES

source

1. 93398  
 Location/Qualifiers  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="genomic DNA"  
 /variety="japonicus"  
 /db\_xref="taxon:34305"  
 /chromosome="6"  
 /clone\_lib="JIT library"  
 /note="TM0082b, a part of TAC clone: TM0082."  
 synonym: Lotus japonicus"

## ORIGIN

## Alignment Scores:

Pred. No.: 6.89e-67 Length: 93398  
 Score: 915.00 Matches: 406  
 Percent Similarity: 17.05% Conservative: 52  
 Best Local Similarity: 15.12% Mismatches: 73  
 Query Match: 30.86% Indels: 2161  
 DB: 15 Gaps: 15

US-09-651-651-5 (1-546) x AP06356 (1-93398)

```

Qy 3 AlaasnseryserserValThrAlaserPheThrValIleAlaValPhePheLeuIleCys 22
Db 34375 TCCAAACAACCGCGCTGCTATGCTTCTAACCGCTCTCGCCGCC-----GTC 34328
Qy 23 GlyValThrThrAlaValAlaGluAspGluThrGluPheHisGlyAspTyrSerLysLeuSer 42
Db 34327 GCCGAGCGCGAGCATGTCGAGACTC-----GACTACTCGAAGCTCTCC 34283
Qy 43 GlyIleIleIlePheProGlyPheAlaserThrGluLeuArgAlaTyrPseIleLeuAspCys 62
Db 34282 GGCATTATCATCCCGGCTTGCTGCTCCACTGAGCTCGAGCATGTCGATTTCTTGACTGC 34223
Qy 63 ProTyrPheProLeuAspPheAsnProLeuAspLeuValTyrLeuAspThrThrLysLeu 82
Db 34222 CTTACTCTCTCTCGATTTCACCTCTGATTTGGTGTGGTGGACACCAACAAAGT- 34164
Qy 82 ----- 82
Db 34163 ACGTTTCACGCAATCTTCTGTGTTCAATCTTCTTCTAGTGAGATTTTCTTGA 34104
Qy 82 ----- 82
Db 34103 TGATTGTTCACTATTAATACTAATAGTACAGCTTGATTAATTATGAAGAGCTTAGAT 34044
Qy 82 ----- 82
Db 34043 AAGGTTTTTATAGCTTATTTTATAGTAATCTTTCTGAAAGTTAGTTAGAAATAATGTT 33984
Qy 82 ----- 82
Db 33983 CATAGAAGAGAGATACAAATCTCGATGAAGACACTTGTGGAGAGACATGCA 33924
Qy 82 ----- 82
Db 33923 ACTGTTTCTGAAACAGTGACAGAGTCTGCAAAAGAAATAAAGAAAGAGATGTTCAATA 33864
Qy 82 ----- 82
Db 33863 CTACTATAGCTATTTTTTGACTTATTTGAATGAGTTTATGAATTTTATCTAATACAC 33804
Qy 82 ----- 82
Db 33803 TTCTATTGAATTTTGGAACCTTAGCTTACATCCATGAACACTTAGAGATTTGGTTT 33744
Qy 83 -----LeuSerAlaValAsnCysTyrPheLysCysMetValLeuAspProTyr 98
Db 33743 GTTTTGACGCTTCTTCTGCGGTGAATGTGGCTTAGTGCAAGGAGTGCTGATCCTTAC 33684

```

```

Qy 99 AaenGIInThAspHisProGluCysLysSerArgProAspSeriGlyLeuSerAlaIleThr 118
Db 33683 AACCAAGCGAATCACCGTATGCAAGTCTCCCGTGAATGAGTCTTCTGCTATTTACA 33624
Qy 119 GluLeuAspProGlyTyr----- 124
Db 33623 GAACCTGATCTCGTTATATTAACAGGTGTTTGTTTATATCATGTCGTTTGTA 33564
Qy 124 ----- 124
Db 33563 AATATGCTTGTGTTTGTATCATTTGTGATGTTTGTGATTTTGCATGATTTGAGAT 33504
Qy 125 -----11e-ThGlyProLeuSerThrValTyrPysGluTyrLeuLysTyrCys 140
Db 33503 GTAAATTTGTTATTTGACAGAGACTCTTCTGTCGATGTAAGAAAGATTAAGTGTGT 33444
Qy 140 sValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAspTyrPargLeuSe 160
Db 33443 CATTTGAATTTGGATTAAGAGCTTATCATTAATATGCTGTCCATATGATTTGAGACTGTC 33384
Qy 160 rProThrLysLeuGluGluArgAspLeuTyrPheHisLysLeuLys----- 175
Db 33383 ACCGTCAATGCTTGAAGAAAGAGATCTTACTTCATAAGCT- AAGTATGTCATACATCT 33325
Qy 175 ----- 175
Db 33324 TTGCATGTAGTCACACTGACTGATAGTCAGAGACTATCTCTGCTATGATTTGGATAGTA 33265
Qy 175 ----- 175
Db 33264 TCCTTGAATTTGTTACAGTTTGACAGTGGAAATTTCAAGTTGGGCTTGCAATGCTAACGC 33205
Qy 175 ----- 175
Db 33204 TAACTAGAATTTTATCGAATAGTAAACAATGTTTTCATTTTACTGATATGATGTTTC 33145
Qy 175 ----- 175
Db 33144 TGGTAAGAGCAATTTGTTTCTGCTAAACAATTAATCATATGTCATGATTAATGACAA 33085
Qy 175 ----- 175
Db 33084 TCTAAGCCATGAGAAACAAAAATGAGTGTTCGAGGGGAACAATCTCTCCCACTTTT 33025
Qy 175 ----- 175
Db 33024 CTTCTCTTCATTAATTTTGTATCTCTTAACAAGAGATGTTCCATCCACCCCTTCAT 32965
Qy 175 ----- 175
Db 32964 TACCCCTCATTAATTTTACTGCAACAAACATATGACAGACTGCGATTAATAATTGTT 32905
Qy 175 ----- 175
Db 32904 CATTAGAGTCTACTATATCTGAACCTAATAGATTAATAGTATTTCTTATGACTTCTGT 32845
Qy 175 ----- 175
Db 32844 TTCTTCTTCAAGTGTCTTGACAACTGTGTTGTTGTTATTTCAATAGTCATATGATAC 32785
Qy 175 ----- 175
Db 32784 ATAACTAGTACAAATTTTAAAGACAGAAATTTGAGCGCTGTAATATGATATTTGATA 32725
Qy 175 ----- 175
Db 32724 AATATTGCTTAACAATAATGATTAAGTTCTAATAGTAATGTAATGATTCACATATT 32665
Qy 175 ----- 175
Db 32664 TCAAAGTGGAGGTAAATTTTTTCTCAAAATGTCAAGGATTAGATGAATTTTTTCTG 32605

```

QY 176 -----LeuTh 177  
Db 32604 TTAATTCCTTGAGATTTTACATTTCCTCCAAATTGATTCTGTTTGAGATGTAC 32545  
QY 177 rPheGluThrAlaLeuLeuArgLyGlyProSerIleValPheAlaHisSerMetG 197  
Db 32544 ATTGAACGTGTTTAAACCTTCGAGGTGGCCCTCTTAGTTTTGCCCATTCATTTGGG 32485  
QY 197 yAenAenValPheArgTyPheLeuGluTrpLeuArgLeuGluIleAlaProLyHisTy 217  
Db 32484 TAAATCACTTTTTCGTTATTCTTGAGGTGGTGAAGATTAAGCAACCAAAACATTA 32425  
QY 217 rLeuSerTrpLeuSerGlnHisIleHisAlaTyPheAlaVal----- 231  
Db 32424 TTTGCAATGGCTGATCAACATATATCAGCATATTTTGCTGT-TGGTATGTTTACATATG 32366  
QY 231 ----- 231  
Db 32365 TGCATTATTACATTTTATCAGATTCGTGATTAGTATATATGTTGCTGATTTTC 32306  
QY 232 -----GlyAlaProLeuLeuGlySerValGluAlaIleLySerTh 245  
Db 32305 CTTTTCATCGGTGTCTCAGAGACTCCCTTTTGGGTGCACAGAAACCGTGAAAGCAAC 32246  
QY 245 rLeuSerGlyValThrPheGlyLeuProValSerGlu----- 257  
Db 32245 ACTCTCTGGTTTCACATTTGGTCTTCATATCTGAGTGAACTTTTAATATTTGTAAT 32186  
QY 257 ----- 257  
Db 32185 TAGATTTTCTTTTCGGTTTGCCTTTTATCTGACATTATTAAGTTTACTGCTGTACA 32126  
QY 257 ----- 257  
Db 32125 AAATATAGGTCTCTAATTCCTTCAGATATTCGCAATTCACAGTTTATTTTAAATGGCT 32066  
QY 257 ----- 257  
Db 32065 TTAAGTACTTCTTGCTTCATTAATTGATTAATTTGATATGACATAATATGTA 32006  
QY 257 ----- 257  
Db 32005 CTGGTACTCTCTCTTCTTCAGACAGAACCTGGAAATTCCTATTTGTTCCAGACTTAA 31946  
QY 257 ----- 257  
Db 31945 AACCAATTAAAGTTAAATTATTAATAAAAACTTATATTAATTGTAATGGCCCAAGCTT 31886  
QY 257 ----- 257  
Db 31885 GCCTAGGACATATTGAATTACACAAACCAACACAGAGGTGAATTTGATTTAAA 31826  
QY 257 ----- 257  
Db 31825 TGAATAGTTGACTACACAACTAGCAACATACACACTTACCAGTGTATAGACATG 31766  
QY 257 ----- 257  
Db 31765 CTTTCTGAAGTCTGTATTCCTCAGAAAGTCTTTAAAGACTTACGATTAAGCTTTCT 31706  
QY 257 ----- 257  
Db 31705 GTTTTATGGTTTGCATTTTCAGTTGTCAACATGCAATTAATAGGTATTCCAAGTAA 31646  
QY 257 ----- 257  
Db 31645 GACTATAAAGTTGAAAAAGTAAACAGGATTAACACACTTGAATTCGCATCAAGAAATG 31586  
QY 257 ----- 257  
Db 31585 TTGATCCAATATTAATCTGTCTAATGCTGTAAATGGCGATGACAGACATGGCTAGAGC 31526  
QY 257 ----- 257

Db 31525 CAAAACCTGTATATATGGATTGGACAAATATATGTTGGGATTAATGCTGTATATCGGCC 31466  
QY 257 ----- 257  
Db 31465 ATGGCGAGAGCCAAAAACCATGATGAGATTGACGGCGCATGGCAAAATATTTGATTA 31406  
QY 257 ----- 257  
Db 31405 CACAGCTTAACAGTCATATTAACCCACTTACATGATTTGTGACTTTAATAGATGATGAA 31346  
QY 257 ----- 257  
Db 31345 CATCTCTGTGCTGTGTCTTAACAGGATAGACACCACTTCATTAATTTTATATCT 31286  
QY 257 ----- 257  
Db 31285 TATTTAATCATTAACCTTTGACACCCAGGCAATTCGCTGAGCTGTGAATTAATTCAGCTGT 31226  
QY 257 ----- 257  
Db 31225 TTTTGTCTAACAGCATATCAAACTTCCTGTCTATTATTCATTAGTGTCTTTAATCAT 31166  
QY 257 ----- 257  
Db 31165 CTCTATAGTCATTTGACAGCATATAGAGCTGTAACTCTGAATGTATCCCTGTGGCA 31106  
QY 257 ----- 257  
Db 31105 TGTTCCTAAATATTCACCTATATCATGATCATGATATGGGGGAGATTGGAAACATTTG 31046  
QY 257 ----- 257  
Db 31045 CCTACTCATGATTAAGAGTGGCTTAATGTCCATATTTGGAAAAATCACTTGTCTATGC 30986  
QY 257 ----- 257  
Db 30985 TTGAAGTGGGCTTAATGTCTAGATTTAGATTTAGATTGAATPACAATTTGAGAAAG 30926  
QY 257 ----- 257  
Db 30925 AATATTTCAATTTTAATTTAAACAGAGTTAAATTCAAAGAACAACTTCGAGTAAGG 30866  
QY 257 ----- 257  
Db 30865 GAGAAACGAGAAAGAAATTTTCATCTTGATGAAGTCTCAATATCATCTAGGGCATGGG 30806  
QY 257 ----- 257  
Db 30805 AAATTGAGCAATTTACTATGAATTAATTTCAATGAAGGTAAATTTCTGAAAGCTGCT 30746  
QY 257 ----- 257  
Db 30745 CAATGAAAGAGAGAGAAAAAAGAGACTTCGAACTTTGAAACCTTACATATATAT 30686  
QY 257 ----- 257  
Db 30685 AATTCATTAACCACTATTCCTTTATPAAGTTTTTTTAAAGAAATTAATAAATCCCTTGAGA 30626  
QY 257 ----- 257  
Db 30625 TTTTGTATGTCTTCAATTACATGCTGAAGCTGTGAAATACATCTTTTGTGGAAT 30566  
QY 257 ----- 257  
Db 30565 CCATCTCTCGTTATATATATACATGCTTCATAGTCGTTTCTTATTTGACGATGAAC 30506  
QY 257 ----- 257  
Db 30505 CTCTTAATTAATTCATGATATGTAAGAAATCTTATCTGAAGAGCTAGCTGATACTTA 30446  
QY 257 ----- 257

Db 30445 TAGAAATAGTAGTGAATTCACAGTTGCTGCATCTAGATTATAGAAATTAATTGA 30386  
Qy 257 ----- 257  
Db 30385 AGTTGACACGATTATTGTATCGCCCTCCAAACCTAAATCATATCAACTTCAGATAAAC 30326  
Qy 257 ----- 257  
Db 30325 ATCAAACTGCCATTGTTAATAGATTAAATGATTGGACATGCGTTTATGTAATATT 30266  
Qy 257 ----- 257  
Db 30265 TCATCTGACAGTTCTATTCCTGCTTAATAAATGCTTATCATATAATTTTGGAGTATCA 30206  
Qy 258 -G|YThra|Aa|g|leu|e|u|e|s|e|r|Phe|a|s|e|r|Leu|T|p|e|u|e|t|P|ro|p|h|e| 277  
Db 30205 GGGAAAGCTCGATTGATGTTCAACTCATTTGGTTCATCTATGGATGATGATGCTTTTTC 30146  
Qy 277 r|y|s|a|n|C|y|e|l|y|a|s|p|a|n|t|S|e|r|T|p|T|h|i|s|P|h|e|s|e|r|G|y|a|l|a|l|a|y|b|y| 237  
Db 30145 CAAGTACTGCAGAACAGATATAATATTTGMAACATTTTGCC---GGAGAAAGGCATGC 30089  
Qy 297 s|a|p|l|y|s|a|r|y|a|l|T|h|i|s|C|y|s|a|p|g|u|g|u|T|y|c|l|n|S|e|r|T|y|s|e|r|G|y|T|p|r| 317  
Db 30088 AGGTCATCATATCATATGTTGTGTGATGAGCGAGAAATTCATTCAACTTATCTGATGGCC 30029  
Qy 317 o|T|h|a|n|l|e|l|e|a|n|l|e|g|u|l|e|P|r|o|S|e|r|T|h|r|----- 328  
Db 30028 AACAAAGATATATCAACTTAAATCCCTTCAAG-TCGTGTGAGATTGAGATTCTCTTG 29970  
Qy 328 ----- 328  
Db 29969 CATTTGCTAGAGCAATCTTGTGATGTAAGTGTGACCACTGAGTACTAATAGC 29910  
Qy 329 -----S|e|r|V|a|l|T|h|r|G| 332  
Db 29909 TCCCAACCTGCTAATTTGTGTCTTGATGAGATTTGATGATATCTTCATTTTGA 29850  
Qy 332 u|T|h|a|l|a|e|u|a|l|a|n|e|t|T|h|s|e|r|M|e|t|G|l|u|C|y|e|l|y|e|u|P|r|o|T|h|r|e|u|e|u|s|e|r|P|h|e|t| 352  
Db 29849 GATACCTGATGCCAACTTGTTCACAGATGGAATGGAATGGAATGGAATGGAATGGAATG 29790  
Qy 352 r|a|a|r|g|u|l|e|u|l|a|s|p|g|l|y|h|r|e|u|p|h|e|l|y|a|l|l|e|g|u|a|s|p|T|y|r|a|s|p|o|a|s|e| 372  
Db 29789 AGCTCGGAATATGCTGATGCTCTCTTTTCAAGCAATTGAAGATTATGACCCAGATAG 29730  
Qy 372 r|y|s|a|r|g|e|l|e|u|h|i|e|g|l|e|u|l|y|b|y|b|e|u|----- 382  
Db 29729 CAAGAGCTTTTGTATACAGATTAGAAAGTG-AGTGTATCTCAACAAATATCATCAAAAT 29671  
Qy 382 ----- 382  
Db 29670 TATATTTGCTACAATGTTTTCATCATTTTACATGAGATGCCGAAATGAGCTTCGAA 29611  
Qy 382 ----- 382  
Db 29610 CTGTAAATAGGCGCTTTCATTTTCAAGGAATATATATTAAGAAACATATGTTCT 29551  
Qy 382 ----- 382  
Db 29550 CTATCTATCTGAAATATGTTTGTGAAAAATAGGTTTTCGAGAGCAAGTTTCTGTTTC 29491  
Qy 382 ----- 382  
Db 29490 TATATAAACTCTATATTTTATTTACTTCTTCCTGATATATACATTTTACTCTGTCG 29431  
Qy 382 ----- 382  
Db 29430 GCAACTTATTTTGTATGTAGTCTACTATTGAACAACATATGTTGAATGCTTTA 29371  
Qy 383 -----T|h|r|i|s|a|s|p|a|r|P|r|o|v|a|l|P|h|e|a|s|P|r|o|l|e|u|h|r|P|r|o|T|p|g|l|u|a|g|P|r| 398  
Db 29370 ATGTTTCAGATCATATCTTTAGTATCTGTTCTTATATCTCACTTACCTGAGACCGGCC 29311

Qy 398 o|P|r|o|l|e|y|s|a|n|v|a|l|P|h|e|C|y|e|l|l|e|T|y|r|G|l|y|----- 408  
Db 29310 ACCAATTAATAAATGCTCTCTGATTTATGCACTGATTCAGAACAAAGGTATAGTTTCC 29251  
Qy 408 ----- 408  
Db 29250 CTCTACAGTGAACATGCTGATGTTTCTATATGTCACCTCAATGCAATGCAAGAAAT 29191  
Qy 408 ----- 408  
Db 29190 TCCATTTAATGAATTTGGTATTATTACTGAGATCATGAACCTGGATTTGGAAATATGA 29131  
Qy 408 ----- 408  
Db 29130 AGTATCAAACTGTGAACACTGAGTGAATTCATCTCCATCATCTGTAAGTTGTTT 29071  
Qy 408 ----- 408  
Db 29070 GAGTTTGAACCTCCATCACAATTTATATTAGTTTACTATGTCGACATTGCATTTTTTT 29011  
Qy 408 ----- 408  
Db 29010 TTGAAAAAGCAAGATATTGGGCTTTATATGAGGCACTACATATGTGAACAATGAT 28951  
Qy 408 ----- 408  
Db 28950 AATTCGTTTCAAGCTGAGAAATGCCATCTTTTATGTTTATGAAATATTATATATT 28891  
Qy 409 -----A|a|h|i|s|e|l|e|u|l|y|s|-T|h|r|----- 413  
Db 28890 GTTGAAGTCGACAGAGACTTGTAGATTGTTTGATTTGATTCATTTATACTACCAT 28831  
Qy 413 ----- 413  
Db 28830 CAATTTTATTTGCAACACTTAATAATTCCTTCCCTGGAGGTAGTTGTGTTAACA 28771  
Qy 413 ----- 413  
Db 28770 AMCCACAGCCACATGATATGATATACAACTAAGAAATTTTCATACGAGTAGAATA 28711  
Qy 413 ----- 413  
Db 28710 AAAAGATTATCTGCTCTCTTGTAGTTCAATTTTAAAGTTATATAGAACTTTCAAAACA 28651  
Qy 413 ----- 413  
Db 28650 TTCAATTTCTTGTCTCTCTATCTTTTAGCAGAAATTCACATTTTAGAGACCTTTT 28591  
Qy 413 ----- 413  
Db 28590 ACATGTGATATATTCGCGTCTCTTCTTATTTGTGAGTGTGCTCATTTGTGAGTACGA 28531  
Qy 413 ----- 413  
Db 28530 ACAATTTCAAGGTATAGGGTAAACATGCTTTACTGACCTTGAATGCTTGACGCTTG 28471  
Qy 413 ----- 413  
Db 28470 CAGCTAGCATAGCTAGAAACATTTGAATGTTTTAGCTTCATGAATAATGCAATATACAA 28411  
Qy 413 ----- 413  
Db 28410 AATCTTTTGTGTAACAATTTTTCCTGACTGTGATCATTTTTCCTGTACATGAGATT 28351  
Qy 413 ----- 413  
Db 28350 CTTAATTTATCTCAGTAAGTCTTTTGGGAAATTTGTTATATCAGAAATAGTATATGA 28291  
Qy 414 -----g|u|a|l|G|l|y|T|h|r|P|h|e|a| 420  
Db 28290 CTTATTTGGAATGATATACATATAGATCTTAACTATCAAAAACAGGTTGGTTACTACTTGG 28231



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 22, 2005, 03:07:09 ; Search time 635 Seconds  
(without alignments)  
5730.586 Million cell updates/sec

Title: US-09-651-651-5  
Perfect score: 2965  
Sequence: 1 MGANSKSVTASFTVIAVFL.....DSESIPIKRTAVWELDKSGY 546

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n.model -DEV=xlh  
-Q/cgnt2/USFTO.spool/US09651651/runat\_18112005\_101830\_20237/app\_query.fasta\_1.711  
-DB=N-geneseq -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LDOFCL=0 -LDOFEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=D10sum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09651651 @CGN 1.1 542 @runat\_18112005\_101830\_20237 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 21:\*

1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2965	100.0	1641	5	AAS01082 Arabidops
2	2944	99.3	1902	12	ADFA7823 Arabidops
3	2332	78.7	3896	3	AAC64438 Arabidops
4	2310	77.9	3896	3	AAC64448 Arabidops

5	444	15.0	328	5	AAS01096	Aas01096 Corn ster
6	345	11.6	1872	12	ADFA7831	Adf47831 Schizosac
7	339	11.4	1893	13	ADT99409	Adt99409 Fusarium
8	325.5	11.0	2326	14	ADW20631	Adw20631 Y. lipoly
9	300.5	10.1	2004	12	ADFA7835	Adf47835 Crepis pa
10	299.5	10.1	1701	12	ADFA7817	Adf47817 Yeast mem
11	299.5	10.1	1986	3	AAC64431	Aac64431 Saccharom
12	299.5	10.1	1986	3	AAC64440	Aac64440 Saccharom
13	299.5	10.1	1986	5	AAS01342	Aas01342 Yeast LCA
14	299.5	10.1	1986	12	ADFA7816	Adf47816 Yeast acy
15	299.5	10.1	1986	13	ADT87255	Adt87255 Yeast Str
16	291.5	9.7	356	5	AAC64441	Aac64441 Saccharom
17	288	9.6	2398	9	AAL62909	Aal62909 Sunflower
18	284	9.6	2700	9	AAL62908	Aal62908 Soybean c
19	283.5	9.5	2030	9	AAL62910	Aal62910 Wheat cDN
20	281	9.5	2479	9	AAL62905	Aal62905 Guayule c
21	280	9.4	2312	3	AAC64432	Aac64432 Schizosac
22	278	9.2	2016	12	ADFA7833	Adf47833 Crepis pa
23	273	9.1	2312	3	AAC64442	Aac64442 Schizosac
24	271	9.1	4093	9	AAL62913	Aal62913 Soybean c
25	269	8.9	2223	12	ADFA7830	Adf47830 Aspergill
26	265	8.9	709	3	AAC64449	Aac64449 Lycopersi
27	263	8.9	2433	9	AAL62907	Aal62907 Rice cDNA
28	263	8.9	709	3	AAC64439	Aac64439 Lycopersi
29	261.5	8.8	7299	13	ADFA7833	Adf47833 Crepis pa
30	261	8.8	2433	9	AAL62907	Aal62907 Rice cDNA
31	259.5	8.8	2427	3	AAC64434	Aac64434 Arabidops
32	257.5	8.7	2565	9	AAL62906	Aal62906 Corn cDNA
33	256	8.6	2047	12	ADFA7829	Adf47829 Aspergill
34	255.5	8.6	1803	12	ADFA7821	Adf47821 Arabidops
35	255.5	8.6	2013	6	AAS01341	Aas01341 Arabidops
36	255.5	8.6	2016	6	ADG88186	Adg88186 A. thalia
37	255.5	8.6	2016	12	ADFA7819	Adf47819 Arabidops
38	255.5	8.6	2425	10	ADG25155	Adg25155 P. patens
39	234	7.9	1137	2	AAV68563	Aav68563 Murine le
40	234	7.9	1236	8	ABZ59338	Abz59338 Mouse LCA
41	234	7.9	1236	10	ADZ25554	Adz25554 Murine LCP
42	234	7.9	1236	10	ADG14388	Adg14388 Murine LL
43	232	7.8	1998	8	ABZ76360	Abz76360 A. thalia
44	232	7.8	1998	12	ADFA7827	Adf47827 Arabidops
45	232	7.8	1998	12	ADFA7827	Adf47827 Arabidops

## ALIGNMENTS

RESULT 1	AAS01082	standard; DNA; 1641 BP.
ID	AAS01082	
XX	AAS01082;	
AC		
XX		
DT	31-MAY-2001 (first entry)	
XX		
DE	Arabidopsis thaliana sterol acyltransferase LCAT2 DNA.	
XX		
KW	Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;	
KW	acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;	
KW	nutritional supplement; dairy product; food product; salad dressing; ds.	
XX		
OS	Arabidopsis thaliana.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1641
FT		/*tag= a
FT		/product= "LCAT2"
XX		
PN	W0200116308-A2.	
XX		
PD	08-MAR-2001.	
XX		
PF	30-AUG-2000; 2000WO-US023863.	
XX		
PR	30-AUG-1999; 99US-0152493P.	



XX (MONS ) MONSANTO CO.  
PA  
XX Laesener M, Van Eenennaam A;  
XX WPI: 2001-169010/17.  
XX P-PSDB; AAU00459.  
DR  
XX New isolated nucleic acid encoding plant lecithin:cholesterol  
PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase  
PT -like polypeptides, for modifying the sterol content and oil production  
PT of plants.  
XX  
PS Claim 5; Page 77; 127bp; English.  
XX  
XX The present sequence encodes for Arabidopsis thaliana  
CC lecithin:cholesterol acyltransferase-like 2 (LCAT2). Several novel  
CC polynucleotides encoding the plant sterol acyltransferases LCAT  
CC (AA5001081-AA501104, AA501341) and ACAT (acyl CoA:cholesterol  
CC acyltransferase-like; AA501311-AA501319) are described. A yeast LCAT  
CC related open reading frame, LKO1 gene sequence (AA501342), and a rat ACAT  
CC (AA501105) cDNA sequence are also described. The polynucleotides encoding  
CC LCAT or ACAT are used to produce LCAT or ACAT polypeptides. They can also  
CC be used in a recombinant construct to transform a host cell (preferably  
CC of a plant) or a plant. The recombinant construct is used to increase or  
CC decrease the sterol content of the host cell or plant. It can be used to  
CC alter oil production of the cell or plant, preferably by increasing it.  
CC The oil of the plant or the plant itself is used as a food product, or as  
CC nutritional or dietary supplements, or in pharmaceutical compositions for  
CC lowering cholesterol. The oil can be used in foods e.g. margarine,  
CC butter, cooking oil, and dressings e.g. salad dressings, mayonnaise,  
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,  
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and  
CC beverages. The alteration in sterol content and/or composition can also  
CC provide a plant with tolerance to stress and insect damage  
XX  
SQ Sequence 1641 BP; 438 A; 355 C; 378 G; 470 T; 0 U; 0 Other:  
  
Alignment Scores:  
Pred. No.: 7,46e-288 Length: 1641  
Score: 2965.00 Matches: 546  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0  
  
US-09-651-651-5 (1-546) x AA501082 (1-1641)  
QY 1 MetG1A1AaSer1ySserVal1ThrAlaSerPheThrVal1IleAlaValPhePheLeu 20  
DB 1 ATGGAGCGAATTCGAAATCAGTAACGGCTTCTTCAACGCTCATCGCGTTTTCCTGG 60  
QY 21 IleCyBg1yG1yArGThrAlaVal1G1uAspG1uThrG1uPheH1sg1yAspTyrSer1y 40  
DB 61 ATTGGGGTGGCGGAATCGCGGTGAGAGATGACGACGAGTTTCAACGGCGATCTGAGAG 120  
QY 41 LeuSerG1y1Ile1IlePProG1yPheAlaSerThG1uLeuArgAla1TrpSer1Ileu 60  
DB 121 CTATCGGATATATCATTCCTCGGATTTGCTCGACACGACTACGACG1GTGATCTT 180  
QY 61 AAspCyProG1yThrProLeuAspPheAsnProLeuAspLeuVal1TrpLeuAspTrpTr 80  
DB 181 GATGTCCATACCTCGTGGACTTCATCCGCTGACCTCGATATGGCTAGACACACT 240  
QY 81 LysLeuLeuSer1AlaValAsnCyS1TrpPhe1yCyMeCVal1LeuAspProTyrAaG1n 100  
DB 241 AAGCTTCTTTCGCTGCAACCTGCTGGTTTAAGTATGCTAGATCCTTATATCA 300  
QY 101 ThrAspHis1ProG1uCy1yS1ySer1ArgProAspSerG1yLeuSerAla1IleThrG1uLeu 120  
DB 301 ACAGACCATTCGCCAGTGAAGTACACGCGCTGACAGTGGTCTTTCACGCAATCACAAGAATTG 360  
QY 121 AspProG1yTyr1IleThrG1yProLeuSerThrVal1TrpLysG1uTrpLeuLys1TrpCy 140

DB 361 GATCCAGGTTACATACAGGTCCTCTTCTACTGCTCGGAAGAGTGGCTTAAGTGGTGT 420  
QY 141 ValG1uPheG1y1IleG1uAlaAsnAla1IleValAlaVal1ProTyrAsp1TrpArgLeuSer 160  
DB 421 GTTGAGTTGGTATGGAAGCAAAATGCGCTGCTTCATACGATTTGGAGATTTGCA 480  
QY 161 ProThrLys1LeuG1uArgAspLeuTyrPheH1s1yLeuLysLeuThrPheG1uThr 180  
DB 481 CCAACCAAAATTGGAAGAGCGTGAACCTTACTTTCACAGCTCAAGCTTAACCTTTAAACT 540  
QY 181 AlaLeuLysLeuArgG1yG1yProSer1IleValPheAlaH1sSerMeC1yAsnAsnVal 200  
DB 541 GCTTAAAACTCCGGGGGCCCTCTTATAGATTTGGCCANTCAATGGGTAAATATGTC 600  
QY 201 PheArgTyrPheLeuG1uTrpLeuArgLeuG1u1IleA1ProLysH1sTyrLeuLys1Trp 220  
DB 601 TTCAATGATCTTCTGGAATGGCTGAGGCTAGAAATGCAACCAAAATTTATTTGAAGTGG 660  
QY 221 LeuAspG1uH1s1IleH1sAla1eThrPheAlaVal1G1yAla1ProLeuLeuG1ySerValG1u 240  
DB 661 CTTGATCAGCATATCATGCTTATTTCCGTGTGAGCTCTCTTCTTGTTCTGTGAG 720  
QY 241 Ala1IleLysSerThrLeuSerG1yVal1ThrPheG1yLeuProValSerG1uG1yThrAla 260  
DB 721 GCAATCAATCATCTCTCTGCTGTTAAGTTGGCTTCCTGTTCTGAGGGAATGCTGT 780  
QY 261 ArgLeuLeuSerAsnSerPheAlaSerSerLeu1TrpLeuMetProPheSer1yAsnCy 280  
DB 781 CCGTTGTGTCMAATTCCTTTTGGCGTCATGTGTGGCTTAATGCCATTTTCMAAGAAATGTC 840  
QY 281 LysG1yAspAsnThrSer1TrpThrH1sPheSerG1yG1yAla1yLys1yAspLysArg 300  
DB 841 AAGGGATGATACACATCTTGACGCAATTTTCTGGGGGTCTGCMAAGAAATTAAGCC 900  
QY 301 Val1TyrH1sCyAspG1uG1uG1uTyrG1uSer1yTyrSerG1y1TrpProThrAsn1Ile 320  
DB 901 GATACCACTGATGAGAGAGGAATATCATCAATAATTTCTGGCTGGCGCAATATAT 960  
QY 321 IleAsn1IleG1u1IleProSerThrSerVal1ThrG1uThrAlaLeuValAsnMetThrSer 340  
DB 1081 CTTTCCAAAGCAATGAGAGACTATGACCCAGATTAACAAGAGATGTATACACAGTTAAAG 1140  
QY 381 LysLeuTyrH1sAspAspProVal1PheAsnProLeuThrProTyrG1uArgProPro1Ile 400  
DB 1141 AAGTTGTATCATGATGACCTCTGTTTATCTCTGACCTCTTGGAGAGACACACTTAA 1200  
QY 401 LysAsnVal1PheCy1IleTyrG1yAlaH1s1LeuLysThrG1uVal1G1yTyrPheAla 420  
DB 1201 AAAAATGATTTTGCATATATGCTCATCTCAAAGACAGAGATGGTTATTAACCTTGGC 1260  
QY 421 ProSerG1yLysProTyrProAspAsnTrp1Ile1IleThrAsp1Ile1eTyrG1uThrG1u 440  
DB 1261 CCAAGTGGCAAACTTTATCTGATATGGAATGGAATCAACGGAATATCATTTACGAACCTGAA 1320  
QY 441 G1ySerLeuVal1SerArgSerG1yThrVal1yLysArg1yAsnAlaG1yPro1IleThrG1y 460  
DB 1321 GGTTCCCTCGTGTCAAGGTCTGGAACCTGTGTGTATGGGAAGCTGGACCTTAATCTGGG 1380  
QY 461 AspG1uThrVal1ProThrH1sSerLeuSerTrpCy1yLysAsnTrpLeuG1yProLysVal 480  
DB 1381 GATGAGACGGTACCCATATATCACTCTTGTGTGCMAAATTTGGCTCGAATTAAGTT 1440  
QY 481 Asn1IleThrMetAlaProG1uProG1uH1sAspG1ySerAspVal1H1sVal1G1uLeuAsn 500

DB 1441 AACATAACATGGCTCCAGCCAGAACGATGAGGACGTACATGTGGAATAAAT 1500  
QY 501 VALAAPHHISGLIHISGLISERAEPIIEIIEAASMeThriylA1ProarValIys 520  
DB 1501 GTTGATCAAGAGCATGGCTGACATCATACCTAATACATGACAAACACCAAGGTTAAG 1560  
QY 521 TYTIEThPheTYGluApsSerGluSerIleProGlyLysArgThraIaValITPgiu 540  
DB 1561 TACATAACCTTTATAGAAGCTCTGAGAGCATTCGGGGAGAGAACCGACGTCTGGAG 1620  
QY 541 LeuApsLysSerGlyTyr 546  
DB 1621 CTTGATAAAAGTGGGTAT 1638  
RESULT 2  
ADP47823  
ID ADP47823 standard; DNA; 1902 BP.  
XX  
XX ADP47823;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Arabidopsis thaliana acyltransferase nucleotide sequence SEQ ID NO:8.  
XX  
XX acyltransferase; enzyme; membrane-spanning region;  
KM active membrane independent acyltransferase; fatty acid ester;  
KM fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;  
KM phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
KM acylated amino acid; structured lipid; fat-soluble molecule; gene; db.  
OS Arabidopsis thaliana.  
XX  
PN W02003100044-A1.  
XX  
PD 04-DEC-2003.  
XX  
PF 28-MAY-2003; 2003MO-SE000870.  
XX  
PR 29-MAY-2002; 2002SE-00001581.  
PR 29-MAY-2002; 2002JUS-0383889P.  
PR 20-JAN-2003; 2003SE-00000142.  
XX  
XX (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.  
PI Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;  
DR WPI; 2004-053268/05.  
DR P-PSDB; ADP47824.  
XX  
XX  
PT New nucleotide sequence encoding an improved acyltransferase polypeptide,  
PT useful for producing structured lipids or fat-soluble molecules, in  
PT removing undesirable fat or in modifying lipids in animal or plant raw  
PT material.  
XX  
PS Claim 5; SEQ ID NO 8; 91pp; English.  
XX  
CC The present invention describes a nucleotide sequence (I) derived from a  
CC nucleotide sequence encoding an acyltransferase polypeptide comprising at  
CC least one membrane-spanning region. (II) encodes an improved active  
CC membrane independent acyltransferase polypeptide in which at least one  
CC amino acid residue of the membrane-spanning region has been deleted  
CC and/or substituted as compared to the original acyltransferase  
CC polypeptide, where the encoded active membrane independent  
CC acyltransferase polypeptide can produce fatty acid esters and/or fatty  
CC acid thioesters such as triacylglycerols, diacylglycerols,  
CC monoacylglycerols, phospholipids, glycolipids, lysolipids, waxesters,  
CC acylated carbohydrates and acylated amino acids. Also described: (1) a  
CC nucleotide sequence molecule (II) comprising at least one promoter region  
CC which functions in a host, where the promoter region is operably linked  
CC to at least one (I), which is operably linked to at least one non-  
CC translated region which functions in a host; (2) a vector comprising (II)  
CC / (3) a host cell comprising (II) or the vector; (4) producing an  
CC improved active membrane independent acyltransferase polypeptide,

CC comprising providing the above host cell and a growth medium preparing a  
CC host cell culture, culturing the host cell culture and harvesting the  
CC host cell culture and recovering the improved active membrane independent  
CC acyltransferase polypeptide; (5) a polypeptide obtained by the method in  
CC (4) where the polypeptide at least is an improved active membrane  
CC independent acyltransferase polypeptide; (6) an oligonucleotide  
CC specifically hybridizing to (I) under stringent conditions; and (7) a kit  
CC comprising the above polypeptide and a stabilizer. The nucleic acid  
CC molecule (I) and the polypeptide are useful in producing structured  
CC lipids or fat-soluble molecules, in removing undesirable fat or in  
CC modifying lipids present in animal and plant raw material. The present  
CC sequence is used in the exemplification of the present invention.  
XX  
SQ Sequence 1902 BP; 516 A; 405 C; 444 G; 537 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 1,2e-285 Length: 1902  
Score: 2944.00 Matches: 541  
Percent Similarity: 99.63% Conservative: 3  
Best Local Similarity: 99.08% Mismatches: 2  
Query Match: 99.29% Indels: 0  
DB: 12 Gaps: 0  
US-09-651-651-5 (1-546) x ADP47823 (1-1902)  
QY 1 MetGlyAlaAsnSerLysSerValThraIaSerPheThraIaValPhePheLeu 20  
DB 1 ARGGAGCGCAATTCGAAATCAGTAACGGCTCTTCAACGGTACGCCGCTTTTCTTG 60  
QY 21 IIECYGELGYIARgThraIaValGluApsGluThraIaPheHisGlyApsTyrSerLys 40  
DB 61 ATTTGGCGTGGCGGAAACGGCGGTGAGATGAGACCGAGTTTCACGGCAGTACTGAAG 120  
QY 41 LeuSerGlyIleIleIleIleProGlyPheAlaSerThraIaValArgAlaTrpSerIleLeu 60  
DB 121 CTATCGAGTATATATCTTCGGGATTTGGCTGACGACACTACGACCGGTGATCCTT 180  
QY 61 AspCysProTyrThraIaPheApsPheApsProLeuApsLeuValITrpleuApsThra 80  
DB 181 GACTGCCATACACTCGCTTGGAGCTTCAATCGCTGACCTGATAGGCTAGACACACT 240  
QY 81 LysLeuLeuSerAlaIaAsnCystrPheLysCysMetValLeuApsProTyrArgGln 100  
DB 241 AAGCTTCTTCTGCTGCAACTGCTGTTAAGTATGATGCTGATATCTTATATCA 300  
QY 101 ThrApsHisProGluCysLysSerArgProApsSerGlyLeuSerAlaIleThraGluLeu 120  
DB 301 ACNAGACATCCCGAGGTAAAGTACACGGCTGACAGTGTCTTCAACCCATCAGAAATG 360  
QY 121 AspProGlyTyrIleThraIaPheLeuSerThraIaITrpleuGluITrpleuLysTyrCys 140  
DB 361 GATCCAGGTTACATAACAGTCTCTTTCTACTGTCTGAAAGAGTGGCTTAAGTGTGT 420  
QY 141 ValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrTrpApsTrpApsLeuSer 160  
DB 421 GTTGAGTTTGGTGTAAAGCAATGCAATTTGCGTTCATACAGATTTGAGATTGCA 480  
QY 161 ProThraIaLeuGluGluArgApsLeuTyrPheHisLysLeuLysLeuThraPheGluIThr 180  
DB 481 CCNACCAATTTGGAAGAGCTGACCTTTACTTTCAACAGTCAAGTGAACCTTTGAAGT 540  
QY 181 AlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHisSerMetGlyApsApsVal 200  
DB 541 GCTTTAAACTCGTGGCGGCCCTTCTATAGTATTTGCCATTCATATGGGTAAATATGTC 600  
QY 201 PheArgTyrPheLeuGluITrpleuApsGluGluIleAlaProLysHisTyrLeuLysITrP 220  
DB 601 TTGAGTACTCTTGTGAAGTCTGAGGCTGAAGATTGCAACAAACATTAATTTGAAGTGG 660  
QY 221 LeuApsGlnHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuGlySerValGlu 240  
DB 661 CTTGATCAGCATATCCATGCTTATTTTCGCTGTTGAGGCTCCTCTTGTGTTGTGAG 720

```

QY 241 AAlaIleYseSerThrLeuSerGlyValThrPheGlyLeuProValSerGlyThrAla 260
DB 721 GGAATCAATCTACTCTCTGTGTGAAGTTGGCTTCTGTTTCTGAGGAACCTGCT 780
QY 261 ArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCyS 280
DB 781 CGGTTGTTGCAATCTTCTTGGCTGTGATCTGTGGCTTATGCATTTTCAAGAATTCG 840
QY 281 LysGlyAspAsnThrSerTrpThrHisPheSerGlyValAlaLysAspLysArg 300
DB 841 AAGGGTGATTAACACTTCTGAGACGCAATTTTCTGGGGGGGCTGCACAAAGAAATAGCCG 900
QY 301 ValTrpHisCysAspGluGluGluTrpGlnSerLysTrpSerGlyTrpProThrAsnIle 320
DB 901 GATATCCACTGTGATGAGAGAAATATTCATCAAAATATTCGTGGCTGGCCCAAAATATT 960
QY 321 IleAsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsnMetThrSer 340
DB 961 ATTAACTTGAATTCCTTCCACTAGCGTTACGAAACAGCTCTAGTCAACATGACCAACG 1020
QY 341 MetGluCysGlyLeuProThrTrpLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThr 360
DB 1021 ATGGAAATGGGCTCCCAACCTTTGTCTTTCACAGCCCGTGAATACAGATGGGACT 1080
QY 361 LeuPheLysAlaIleGluAspTrpAspProAspSerLysArgMetLeuHisGlnLeuLys 380
DB 1081 CTTTTCAAAGCAATAGAAACATATGACCCAGATAGCAAGATGTTACCAAGTTAAG 1140
QY 381 LysLeuTrpHisAspAspProValPheAsnProLeuThrProTrpGluArgProIle 400
DB 1141 AAGTTGTATCATGATGATGACCTGTTTAACTCTGACTGCTGGAGAGACACACTATA 1200
QY 401 LysAsnValPheCysIleTrpGlyValAlaIleLeuLysThrGluValGlyTrpTrpPheAla 420
DB 1201 AAAAAGTATTTTGGCATATATGCTGTGCTCATTAACAGAGAGGTGTTATTAATCTTGGC 1260
QY 421 ProSerGlyLysProTrpProAspAsnTrpIleIleThrAspIleIleTrpGluThrGlu 440
DB 1261 CCAAGTGGCAACCTTATCTCTGATTAATGATGATCATACCGATATCTTTTAAACTGAA 1320
QY 441 GlySerLeuValSerArgSerGlyThrValValAspGlyAsnAlaGlyProIleThrGly 460
DB 1321 GGTTCCCTCGTGTCAAGTCTGGAACGTGTGATGAGGAACGCTGGAACCTTAACCTGGG 1380
QY 461 AspGluThrValProTrpHisSerLeuSerTrpCysLysAsnTrpLeuGlyProLysVal 480
DB 1381 GATGAGACCGTACCCATATTCACCTCTGTGTGCAAAATTTGGCTCGACCTAAAGTT 1440
QY 481 AsnIleThrMetAlaProGlnProGluHisAspGlySerAspValHisValGluLeuAsn 500
DB 1441 AACATTAACAATGGCTCCCAAGCCAGAACAGATGAAAGCAAGCTTACATGTGAACCTAAAT 1500
QY 501 ValAspHisGluHisGlySerAspIleIleAlaAsnMetThrLysAlaProArgValLys 520
DB 1501 GTTGTATCATGAGCATGGGTCAAGACATCATAGCTAAATGATCAAAACACCAAGGGTTAAG 1560
QY 521 TrpIleThrPheTrpCysLysAspSerGluSerIleProGlyLysArgThrAlaValTrpGlu 540
DB 1561 TACATTAACCTTTTATATAAGCTCTGAGAGCATTTCCGGGGAAGAAACCGACGTCTGGAG 1620
QY 541 LeuAspLysSerGlyTyr 546
DB 1621 CTTGATTAACAACAAATCAC 1638

```

```

RESULT 3
AAC64438
ID AAC64438 standard; DNA; 3896 BP.
XX
AC AAC64438;
XX
DT 12-FEB-2001 (first entry)
XX
DE Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:11.

```

```

XX PDAT: phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KM EST: expressed sequence tag; fatty acid; oil content; ds.
XX Arabidopsis thaliana.
XX WO20060095-A2.
XX 12-OCT-2000.
XX
XX 28-MAR-2000; 2000WO-EP002701.
XX
XX 01-APR-1999; 99EP-00106656.
XX 10-JUN-1999; 99EP-00111321.
XX 07-FEB-2000; 2000US-0180687P.
XX
XX (BADI ) BASF PLANT SCI GMBH.
XX
XX Dahlqvist A, Stahl U, Lenman M, Banae A, Ronne H, Stymer S;
XX WPI; 2000-665012/64.
XX
XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
XX pathway for triacylglycerol production and DNA encoding them, useful for
XX producing triacylglycerol, or for transforming any cell or organism to
XX increase oil content.
XX
XX Claim 6; Page 60-61; 97pp; English.
XX
XX The present invention describes an enzyme for catalyzing (in an acyl-CoA-
XX independent reaction) the transfer of fatty acids from phospholipids to
XX diacylglycerol in the biosynthetic pathway for the production of
XX triacylglycerol (TAG). The enzyme is designated as
XX phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the
XX nucleotides encoding them are useful for producing TAG and/or TAG with
XX uncommon fatty acids. The enzyme and the nucleotide are also useful for
XX transforming any cell or organism in order to be expressed in this cell
XX or organism and result in an altered, preferably increased oil content of
XX this cell or organism. The present sequence represents the Arabidopsis
XX thaliana PDAT genomic DNA
XX
XX Sequence 3896 BP; 1042 A; 725 C; 769 G; 1360 T; 0 U; 0 Other:
XX
XX
XX Alignment Score:
XX Pred. No.: 1,33e-223 Length: 3896
XX Score: 2332.00 Matches: 541
XX Percent Similarity: 41.58% Conservative: 0
XX Best Local Similarity: 41.58% Mismatches: 5
XX Query Match: 78.65% Indels: 759
XX DB: 3 Gaps: 11
XX
XX US-09-651-651-5 (1-546) x AAC64438 (1-3896)

```

```

QY 1 MetGlyAlaAsnSerLysSerValThrAlaSerPheThrValIleAlaValPhePheLeu 20
DB 1 ATGGAGCGCAATTTGAAATACAGTAAACGGCTTCTTCAACCGTCATCGCGGTTTCTTGG 60
QY 21 IleCysGlyArgThrAlaValGluAspGluTrpGlnPheHisGlyAspTrpSerLys 40
DB 61 ATTTCGGGTGGCCGAACCTGGGTGAGATGAGACCGAGTTTCAAGGCGACTACTCGAAG 120
QY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSerIleLeu 60
DB 121 CTATCGGGTATTAATCATTTCCGGGATTTGCTGTGACGACGCTACGAGCGTGTGATCTT 180
QY 61 AspCysProTrpTrpProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrThr 80
DB 181 GACTGTCCATACACTCCGTTGGACTTCAATCGGCTCGACCTCGTATGGCTAGACACACT 240
QY 81 Lys----- 81
DB 241 AAGGTCCGTGATCTTCAATTTCTTGGCTCTTATTTCTGTGGTGGAGTCACTTGTGATG 300

```

QY 81 ----- 81  
 Db 301 AATTCAGGGAATATAGCAATGAAGCATGTCTCGTCTCTTATGATTCGTCATTA 360  
 QY 81 ----- 81  
 Db 361 GTCACAGTGAAGCGCTTCTGAATCTGAGTTAGATCATATAAACAAGCTGACTGGCGAG 420  
 QY 81 ----- 81  
 Db 421 TGTTCCCATCGCTTTTGTTGCTCGTAATGTAGCGCATGATGTGTAATAGTCTGGC 480  
 QY 81 ----- 81  
 Db 481 TTTTATTAACAAGATCTGCAAGTTTTCAGAGTGTCAATAGTAGTTAGAAAAAGTTA 540  
 QY 81 ----- 81  
 Db 541 GGTCAATTTTACTGTGTGCAATGTGATTCCTTTGGTTGCTTACTGATCGACGTGATGA 600  
 QY 82 ----- 98  
 Db 601 TGCTTTACAGCTCTTTCTCTGTCTGCAACTGCTGGTTTAAGTGTATGCTGCTAGATCTT 660  
 QY 98 TAAAGTATThAspHisProGluCysIysSerArgProAspSerGlyLeuSerAlaIleTh 118  
 Db 661 TAAATCAACAGACCATCCGAGTGTAGTACACGGCTGACAGTGTCTTTCAGGCATCAC 720  
 QY 118 TGIuLeuAspProGlyTyrIleThr ----- 126  
 Db 721 AGAATGTGATCCAGTTACATTAAC-AGGTAGTTCCGATTTTCTTTTGAGTTTC 779  
 QY 126 ----- 126  
 Db 780 TTCAATTTGATATCATCTGTGTGATATAATAGCTAAAGTTCAATTAATTTGGTCAT 839  
 QY 127 ----- 145  
 Db 840 TTTCAGGTCCTCTTCTACTGTCTGGAAGAGTGGCTTAAGTGTGTGAGTTGGTAT 899  
 QY 145 EGIuAlaAsnAlaIleValAlaValAProTyrAspTPrArgLeuSerProThrIysLeu 165  
 Db 900 AGAAGCAAAATGCATTTGTGCTGTTCATAGATGAGATTGTCACCAACCAAAATTTGA 959  
 QY 165 uGlaArgAspLeuTyrPheHisIysLeuLys ----- 175  
 Db 960 AGAGCGTGAACCTTTACTTTTCAAGCTCAAGCTCAAGTATGCTTATCAGGCTAATGTCTT 1019  
 QY 176 ----- 177  
 Db 1020 CTCTCTTTTATGTAAGATGAAGCTAGAGCTGTGTCGTCCTTCTTTTTCAGAGTTGAC 1079  
 QY 177 rPheGluThrAlaLeuIysLeuArgGlyGlyProSerIleValAAlaHisSerMetG 197  
 Db 1080 CTTTGAACACGCTTTTAAACCTCGTGGCGGCTTCTATAGTATTTTGGCCATTAATGGG 1139  
 QY 197 YAsnAsnValPheArgTyrPheLeuGluTPrLeuArgLeuGluIleAlaProLysHisT 217  
 Db 1140 TAAATATGTCTTCAGATACCTTTCTGGAATGSCCTAGAGCTAGAAATTCACCAAAACHT 1199  
 QY 217 rLeuLysTrpLeuAspGlnHisIleHisAlaTyrPheAlaVal ----- 231  
 Db 1200 TTTGAAGTGGCTTATCAGCATATCCATGCTTATTTGCTGT-TGATACGGGCTACTAT 1258  
 QY 231 ----- 231  
 Db 1259 CCTTAAGTACCAATTTATTTTCTCTAATTTGGGGAGTTATGTTGACTTACTGAT 1318  
 QY 232 ----- 241  
 Db 1319 TGAGTCATACCTGATTTGTGTGATTTAGAGACCTCTCTCTGTGGTTTCTGTGAGGC 1378  
 QY 241 aIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGlu ----- 257

Db 1379 AATCAATCTACTCTCTCTGCTGTAACGTTTGACCTTCTCTGTGTTCTGAGGTACCTCTGA 1438  
 QY 257 ----- 257  
 Db 1439 CTTCTCTTAGTTTAAAGTATGATATACACAGGCTTTATTACTCACTGATTTTCC 1498  
 QY 257 ----- 257  
 Db 1499 TTTGAAGTATTAATCTTTTGTAAATGAACCTGTGACGGATATAGTATCTGTAGATCT 1558  
 QY 257 ----- 257  
 Db 1559 GAAAGCTAGTATTAACAAGAACATATTGGGTAGTATACCTGTACGGGCTTACTTAA 1618  
 QY 257 ----- 257  
 Db 1619 TACACCAACACACATGACATGATTTAGTTTTCAGATTATTTATGTAGACTTAAATT 1678  
 QY 257 ----- 257  
 Db 1679 GAGAGAAACCTTGACTGAATCTTTTATTTTAATAGCTATGATTTGTTATTTGAAT 1738  
 QY 258 ----- 260  
 Db 1739 CATGTACATATTGACATGCGCTTCTCATGTTTTTTTGGCAGAGCTTCAGGAACTGC 1798  
 QY 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerIysAsn 280  
 Db 1799 TCGTTGTGTGCCAATCTCTTTTGGTGTGTCATGTTGGCTTATGCCATTTTCAAAAGATT 1858  
 QY 280 sIysGlyAspAsnThrSerTrpThrHisPheSerGlyAlaAlaIleLysAspLysAr 300  
 Db 1859 CAAGGTGATTAACACATTCCTGACGCAATTTTCTGGGGGGCTGCACAAAGAAATAGCG 1918  
 QY 300 gValIYrHisCysAspGluGluGluTyrGlnSerLysIYrSerGlyTyrProThrAsn 320  
 Db 1919 CGTATACCACTGATGAAGAGAAATATCATATAATTTCTGGCTGGCCCAAAATAT 1978  
 QY 320 eIleAsnIleGluIleProSerThrSer ----- 329  
 Db 1979 TATTAACATTGAATTCCTTCCACTAG-CGGTTAGACTGTATATGCAACTGAACACT 2037  
 QY 329 ----- 329  
 Db 2038 AACAAAAGTTTCAACCAAGATGTTCACTCTCATATTTCGTTCCCTTGATGTATCCATC 2097  
 QY 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeu 349  
 Db 2098 AGTTACAGAAACGCTCTTACTCAACATGACACAGATGAATGTGGCTTCCACCTTT 2157  
 QY 349 uSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLysAlaIleGluAspTyrAs 369  
 Db 2158 GTCTTTCACGCCCGTGAATACACAGATGGGACTCTTTTCAAGCAATAGAAAGCTATGA 2217  
 QY 369 pProAspSerLysArgMetLeuHisGlnLeuLysIleu ----- 382  
 Db 2218 CCCAGATAGCAAGAGATGTTTACACACAGTTAAAGAAATA-CGTACCTTCTTTGTGATA 2276  
 QY 382 ----- 382  
 Db 2277 GAATATTTGCTCATTCATCATCTGCTGCTTCTTTGTATACGTCAAAATTTGTTTAA 2336  
 QY 382 ----- 382  
 Db 2337 ATCTATATCAATTTGTTATATGCTTTGTCTTCTTACTATAAGAAACAAGATAATATCA 2396  
 QY 382 ----- 382  
 Db 2397 GAAACCTTATTAATGATTATCAGTTCTCTCTTATATATGAAATGCTTTTTCGTTTAC 2456  
 QY 382 ----- 382

```

Db      2457 AGTTATGATGCAAAAAGGGGATTTTACTGATTGATCTCTCATTTCTACTGTTGTTT 2516
Qy      382 -----
Db      2517 TGACTAATAGCGTCAATTTTGTCTTCTAGCAAAATCTTGTGAATTATATATACATGCT 2576
Qy      383 -----
Db      2577 AACTATACTTTTCAGGTGTATCATGATACACCTGTTTTAACTCTGACTCTCTGGGA 2636
Qy      396 uArpProProlLeuYAsnValPheCysIleTyrgIAlaHisLeuLeuYsThGlu----- 414
Db      2637 GAGACCACTTAATAAAATGTAATTTTGCAATATAGTGCTCATCTAAAGACAGAGGTAG 2696
Qy      414 -----
Db      2697 ATGCATTTCTCAATATCATATTAAGCGTTGACTTTGTTATTAATTCCTCATTTGGTTTGC 2756
Qy      414 -----
Db      2757 AATATCTTTTGAATTATGATTTATCTTCTCCCTGCATCTTATGCTATTAAGCGTTAAA 2816
Qy      415 -----
Db      2817 GGTACTAAATGATATGAAGCTGCTGCATAGGTGGTTATTACTTTGCCCAAGTGGCAA 2876
Qy      424 sProTyPProAspAsnTrpIleIleThrAspIleIleTyrgIuThGluGlySerLeuVa 444
Db      2877 ACCTTATCCCTGAATATGATGATCATCGATATCATTTATGAAGCTGAAGTTCCCTCGT 2936
Qy      444 lser----- 445
Db      2937 GTCAAGGTAATTTTCCGCAATGGCAGAGTAAGAAACAGAAAGCAAGCTTCTGTATCAG 2996
Qy      445 -----
Db      2997 TCTAGTGCATGTTATCTCATGTTGCATTAAGCAAAATTATTAACAACATAAATTATAGTAC 3056
Qy      445 -----
Db      3057 TTTTATCATTCCTTTTGTAGCTTAGTGATGATCAGTGGCTTAAAGTGGAGAGAGTGT 3116
Qy      445 -----
Db      3117 TGCATAAACATGACACTTGATCAAGATTAAGCAAAACAAACTAACCATTTCTG 3176
Qy      445 -----
Db      3177 AATTTCATATTATTAGAGTAGTGTGCTTTTAAAAAATTTGTTTAAAGAACCGAAAAA 3236
Qy      446 -----
Db      3237 CTAGTTCATATCTTGAATTGTGCAATATCTGCAGGTCTGGAACGTGTGATGGGAACGC 3296
Qy      455 aGlyProIleThGlyAspGluThr----- 463
Db      3297 TGGACCTTAATACCTGGGATGAGACGGTAAGCTCAGAACTTGTGTAATTAATCTTCTT 3356
Qy      463 -----
Db      3357 GCMAACTAGAGACTAAGATATACTTGCTTGTGGAACACTGCTTGCTATGTTCTCTA 3416
Qy      464 -----
Db      3417 GTACACTGCAAATATGACTCTCGCTACTTTATTTATGTAATGAATGATCTCTTATAGG 3476
Qy      464 aLProTyRHisSerLeuSerTrpCysLeuAsnTrpLeuGlyProLeuValAsnIleThm 484
Db      3477 TACCTATATCTTACTCTCTTGGTGCAGAAATGGCTCGGACCTTAAGTTAATCAATAACAA 3536
Qy      484 eR1a-Pr----- 486
Db      3537 TGGCTCCCGAGTACTTTTATAGTTCTCACCTTATATAGATCAAACTTAAAGTATAC 3596

```

```

Qy      486 ----- 486
Db      3597 TTTTCTGTATGTGTGATTACTTACCTCCATTGTTCTTTCTTAAATAATCATATATCTCTG 3656
Qy      487 -----
Db      3657 TACTCTCAAGAACTTGTATTAATCTAATACAGAGATTCATTGGGAAAAATAAACACAG 3716
Qy      488 ProGluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySer 507
Db      3717 CCAGAACACGATGGAAGCGAGTACATGTGGAACATAATGTTGATCATGACACATGGGTCA 3776
Qy      508 AspIleIleAlaAsnMetThrIleValAspValArgValIleTyrgIleThrPheTyrgIuAsp 527
Db      3777 GACATCATAGCTTAACATATACAAAGACCAAGGTTAAAGTACATTAACCTTTATGAAAGC 3836
Qy      528 SerGluSerIleProGlyLeuArgThrAlaValTrpGluLeuAspLysSerGlyTyR 546
Db      3837 TCTGAGAGCATTCGGGGAGAGAACCGCAGTCTGGAGCTTGAATAAAAGTGGTAT 3893

```

```

RESULT 4
AAC64448
ID AAC64448 standard; DNA; 3896 BP.
XX
AC AAC64448;
XX
DT 12-FEB-2001 (first entry)
XX
DE Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:10b.
XX
KW PDAT, phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KM EST; expressed sequence tag; fatty acid; oil content; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200060095-A2.
XX
PD 12-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-EP002701.
XX
PR 01-APR-1999; 99EP-00106556.
PR 10-JUN-1999; 99EP-0011321.
PR 07-FEB-2000; 2000US-0180687P.
XX
XX (BADI ) BASF PLANT SCT GMBH.
XX
PI Dahlqvist A, Stahl U, Lenman M, Banae A, Ronne H, Stymne S;
XX WPI; 2000-665012/64.
XX
PT Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DNAs encoding them, useful for
PT producing triacylglycerol, or for transforming any cell or organism to
PT increase oil content.
XX
XX Claim 6; Page 95-96; 97pp; English.
XX
XX The present invention describes an enzyme for catalysing (in an acyl-CoA-
XX independent reaction) the transfer of fatty acids from phospholipids to
XX diacylglycerol in the biosynthetic pathway for the production of
XX triacylglycerol (TAG). The enzyme is designated as
XX phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the
XX nucleotides encoding them are useful for producing TAG and/or TAG with
XX uncommon fatty acids. The enzyme and the nucleotide are also useful for
XX transforming any cell or organism in order to be expressed in this cell
XX or organism and result in an altered, preferably increased oil content of
XX this cell or organism. The present sequence represents the Arabidopsis
XX thaliana PDAT genomic DNA
XX
SQ Sequence 3896 BP; 1042 A; 733 C; 760 G; 1360 T; 0 U; 1 Other;

```

## Alignment Scores:

Pred. No.: 2,18e-221 Length: 3896  
 Score: 2310.00 Matches: 537  
 Percent Similarity: 41.28% Conservative: 0  
 Best Local Similarity: 41.28% Mismatches: 9  
 Query Match: 77.91% Indels: 759  
 DB: 3 Gaps: 11

US-09-651-651-5 (1-546) x AAC64448 (1-3896)

QY 1 MergIyAlaAsnSerlyServalThrAlaSerPheThrValIleAlaValPhePheLeu 20  
 DB 1 ATGGAGCGCAATTCGAATCAGTACGGCTTCCCTACCGCTCATCGCGTTTTCCTG 60  
 QY 21 IlcYgGlyVgIyAqThraIaValGluAspGluThrGluPheHisGlyAspYrSerLyS 40  
 DB 61 ATTTCGGGTGGCGCACTGGCGTGGAGATGAGACCGAGTTTCACGGCGCACTCTGSA 120  
 QY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGluLeuArgAlaTPSerIleLeu 60  
 DB 121 CTATCGGGTAAATCATTTCCCATTTCCGCTCGACGCGACTACGAGCGTGGTGCATCTT 180  
 QY 61 AspCyEProIyThrPProLeuAspPheAsnProLeuAspLeuValTTPLeuAspThrThr 80  
 DB 181 GACTGTCCATACACTCCGTTGCACTTCATCCGCTGACCTCGTATGCTAGACACCACT 240  
 QY 81 Lys----- 81  
 DB 241 AAGGTCGGTATCTTCATTTCCCTTCGCTCTATTCGTGCGTGCAGTCACTTGTGATG 300  
 QY 81 ----- 81  
 DB 301 AATTCAAGGAATATAGCAATGAGCAATGTCGTCTCTATTATTGATTCTGTATTA 360  
 QY 81 ----- 81  
 DB 361 GTCAACAGTGACGCTTGTGAATCTGAAGTTAGATCATATMAACAGCTCAGCGGAG 420  
 QY 81 ----- 81  
 DB 421 TCTTTCCCATCGCTTTTGTTGGTTCGCTAAATGAGCGCAATGAATGTGTAATTAAGTCTGCC 480  
 QY 81 ----- 81  
 DB 481 TTTTATTCACTAGATCTGCAGATTTTTCAGAGTGCATATAGTAGTAAGAAATGTTA 540  
 QY 81 ----- 81  
 DB 541 GGTCATTTTACTTGTGCAATGTGATTCCTTTGGTTGCTTACTGATCGACGTGATGA 600  
 QY 82 -----LeuLeuSerAlaValAsnCyETrpPheLysCySMeValLeuAspProIy 98  
 DB 601 TGGTTTACAGCTTCTTTCTGCTGTCACACTGCTGGTTTAAAGTGAATGGTGTCTAGATCTCTTA 660  
 QY 98 fAsnGlnThzAspHisProGluCySlySerArgProAspSerGlyLeuSerAlaIleTh 118  
 DB 661 TAACTCAACACACATCCCGAGTGAAGTACGGGCTGCAGAGTGGTCTTTTCAGCCATCAC 720  
 QY 118 rGluLeuAspProGlyYrIleThr----- 126  
 DB 721 AGAATGGATCCAGGTTACATTAAC-AGGTAGTTTCGGATTTTCTTTTTCAGATTTC 779  
 QY 126 ----- 126  
 DB 780 TTCAATTTGATATCATCTTGTGTGATATAATAGGCTAAAGTCAATTAATTTGGCAATT 839  
 QY 127 ----GlyProLeuSerThraIITrPlyGluTrpLeuLysTrpCySValGluPheGlyI 145  
 DB 840 TTCAGGTCTCTCTTCTACTGTCTGGAAAGAGTGGCTTAAGTGTGTGTGAGTTGGAT 899  
 QY 145 eGluAlaAsnAlaIleValAlaValIProTyAspTrpArgLeuSerProThrLysLeuG 165  
 DB 900 AGAACCAATGCATTTGTGCTGTTCATACGATTGGAGATTGTCAACCAACCAATTTGGA 959

QY 165 uGluArgAspLeuTyTrpPheHisLysLeuLyS----- 175  
 DB 960 AAGAGCTGACCTTTACTTTCACAGAGCTCAAGTTAGCTTTATCAGGCTAATGCTTTTAT 1019  
 QY 176 -----LeuTh 177  
 DB 1020 CTTCTCTTTTATGTAAGTAAAGCTAAGAGCTGTGGTCTCTTCCTTTTCGAGGTTGAC 1079  
 QY 177 rPheGluThraIaLeuLysLeuArgGlyGlyProSerIleValIleAlaHisSerMetG 197  
 DB 1080 CTTTGAACTGCTTTAAATCTCGTGGCGGCTCTTATAGATATTGCCATTCAAATGAG 1139  
 QY 197 yAsnAsnValPheArgTyPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisTy 217  
 DB 1140 TAAATAAGTCTTCAGATCTTCTGGAATGGCTGAGGCTAGAAATTTGACCAAAACATTA 1199  
 QY 217 rLeuLysTrpLeuAspGlnHisIleHisAlaTyTrpPheAlaVal----- 231  
 DB 1200 TTTGAAAGTGCTGATCAGATATCCATGCTTATTTTCGCTGT-TGGTACCGGCTTACTAT 1258  
 QY 231 ----- 231  
 DB 1259 CCTTAAGTTACATTTTATTTTCTCTAATTGGGGAGTTAATGTGTGACTTATGAT 1318  
 QY 232 ----GlyAlaProLeuLeuGlySerValGluAl 241  
 DB 1319 TGAGCTCGATACCTGATTTTGTGTGATTATTTAGAGGCTCTCTTGTGTGTGTGAGGC 1378  
 QY 241 aIleLysSerThrLeuSerGlyValIthrPheGlyLeuProValSerGlu----- 257  
 DB 1379 AATCAATCTACTCTCTGTGGTGAAGTTCGTTGSCCTTCGTTTCGAGGTACCTCTGA 1438  
 QY 257 ----- 257  
 DB 1439 CTTCTCTTATGTTTAAGTAGTGTGATATCAACAGGCTTATTAATCACTGATTTTCT 1498  
 QY 257 ----- 257  
 DB 1499 TTTGAAAGTATTACTTTTGTTAATGAAGTGTGACGAGATATGATATCTGTAGATCTT 1558  
 QY 257 ----- 257  
 DB 1559 GAAAGTCTAGTATCAAAAGAACATATTGGGATGATACCTGTCAAGCGCTTACGTA 1618  
 QY 257 ----- 257  
 DB 1619 TACAAACCAACACATGTACACTGATTTAGTTTTCAGATTATTATGTAGACTTAAATT 1678  
 QY 257 ----- 257  
 DB 1679 GAGAAACACTTGACTGAATCTTTTATTTAATAGGCTATGATTTGTTATGAAAT 1738  
 QY 258 -----GlyThrAl 260  
 DB 1739 CATGTGACATATTGACATGGGCTTCTCANGTTTTTGTGGCAAGGCTTCAGGAACTGC 1798  
 QY 260 aArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCy 280  
 DB 1799 TCGGTTGTTCCAATCTTTTCGCTGCTGATATGTGGCTTATGCAATTTTCAAAAGATTG 1858  
 QY 280 sLyGlyAspAsnThzSerTrpThrHisPheSerGlyValAlaAlaLysAspLysAr 300  
 DB 1859 CAAGGCTGATACACATTCGTGACGCAATTTTTCGGGGGCTGCACAAAGAAATTAACGC 1918  
 QY 300 gValTyHisCyAspGluGluGluTyArgIleAspLysTySerGlyTyTrpProThrAsnI 320  
 DB 1919 CGATATACACTGTGATGAAGAAATATTCATCAATCAATATTTGCTGGCTGGCCGACAAATAT 1978  
 QY 320 eIleAsnIleGluIleProSerThrSer----- 329  
 DB 1979 TATTATCATTTGAATTCCTTCCACTAG-CGGTTAGACTGTGATATGCAACTGTAACT 2037

```

QY 329 ----- 329
Db 2038 AACAAAAGTTCCACCAAGATGTCCTCTCATATTGCTCTTGATGTGATCCATC 2097
QY 330 -ValThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeu 349
Db 2098 AGTTACAGAAAAGCCTCTAGTCACTGACACAGCATGGAATGGGCTTCCACCTTTT 2157
QY 349 uSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLeuAlaIleGluAspTyrAs 369
Db 2158 GTCTTTACAGCCCGCTGAACCTAGCAGATGGGACTCTTTCAAGCAATGAAAGACTATGA 2217
QY 369 PProAspSerLysArgMetLeuHisGluLeuLysLysLeu----- 382
Db 2218 CCCAGATGACCAAGAGGATGTTACACAGATTAAAGATA-CGTAAGCTTTCTTGATGA 2276
QY 382 ----- 382
Db 2277 GAAATATTGCTCATGCATGATCATCTTGCTGGCTTCTTGACGTCAATTGTTTGTAA 2336
QY 382 ----- 382
Db 2337 ATCTATATCAATGTGTATATGCTTTGCTTTCTTACTATAGAAAGAGATATATCA 2396
QY 382 ----- 382
Db 2397 GAAACCTTATTATTGATTATCACTTCTCTCTTATTATGAAATGCTTTTTCGTTTAC 2456
QY 382 ----- 382
Db 2457 AGTTATGAAATGCAAAAGGGGGATTATTAGTTGATGATTCCTCATTTCTAGTTGTTT 2516
QY 382 ----- 382
Db 2517 TGACTAATAGCGCAATTTTGTGTTTCTAGCAAACTTTGTGTAATTATATATACATGCT 2576
QY 383 -----TyrHisAspAspProValPheAsnProLeuThrProTyrGlu 396
Db 2577 AACCTAATCTTTAGAGTGTGATCATGATGACCTGTTTAACTCTGACTCTTGCGGA 2636
QY 396 uArgProProlIleLysAsnValPheCysIleTyrGlyAlaHisLeuLysThrGlu----- 414
Db 2637 GAAACACACCTATAAAAAAGTATTTTGCATATATGGGCTCATCTAAAGACAGAGATAG 2696
QY 414 ----- 414
Db 2697 ATGCATTCATCATATGACATTATGCGTTGACTTTGTTATTATATCCCAATTTGGTTGC 2756
QY 414 ----- 414
Db 2757 AATATCTTTTGAATTATGATTATCTTCTCCCTTGCACTTATCCATTAAGCGTTAAA 2816
QY 415 -----ValGlyTyrTyrPheAlaProSerGlyLys 424
Db 2817 GGTACTAAATGTAAGAAGTGTCTGTCATAGCTTGGTTATTACTTTGCCCAAGTGCA 2876
QY 424 sProTyrProAspAsnTyrIleIleThrAspIleIleTyrGluThrGluGlySerLeuVal 444
Db 2877 ACCTTATCCGATTAATTGATCATCAGGATATCATTTATGAAGCTGAAGTTCCTCGT 2936
QY 444 lSer----- 445
Db 2937 GTCAAGATTAATTTCCGCAATGCAAGATAAGAAAGCAAGCTTCTGTATCAG 2996
QY 445 ----- 445
Db 2997 TCTAGTGCATGTTATCTCAGTTGATGCAATAGCAAAATTATTAACAACTAAATTAAGTAC 3056
QY 445 ----- 445
Db 3057 TTTTATATCATTCCTTTTAGCTTAGTGATGATGATGCTTAAAGTGGAAGAGGTGT 3116
QY 445 ----- 445

```

```

Db 3117 TGCATGAAACATGACACTTGTATCAAGATACTAGCAAAACAACTAACCCATTCTG 3176
QY 445 ----- 445
Db 3177 AATTTCAATATTATTAGAGTAGTGTGCTTTTAAAAAATTTGTTTAAAGAAACGAAAA 3236
QY 446 -----ArgSerGlyThrValValAspGlyAsnAl 455
Db 3237 CTAGTTCAATCTTGATTGTGCAATATCTGCAGGTCTGGAACCTGTGTTGATGGAAACG 3296
QY 455 sGlyProIleThrGlyAspGluThr----- 463
Db 3297 TGACCTTAATTACTGGGATGAGACCGTAAAGCTCAGAAATTGGTTTGAATTAATCTCTT 3356
QY 463 ----- 463
Db 3357 GCAAACTACTAGAGACTAAGATTAATACCTTGCTTGAACAACCTGCTGATGTTCTGA 3416
QY 464 -----V 464
Db 3417 GTACACTGCATATTGACTCTCCGCTACTTTTATTGATTAGAAATTGATCTTTATAGG 3476
QY 464 aLProTyrHisSerLeuSerTyrCysLysAsnTyrLeuGlyProLysValAsnIleThrM 484
Db 3477 TACCTTATCATTCATCTCTTGCTGCAAGAAATGGCTCGAACCTTAAGTTAACATPACAA 3536
QY 484 eCfAla-Pro----- 486
Db 3537 TGCTCCCGCAGGTACTTTTATTAGTTCTGCACTTAATATGATCAAACTTTAAGTTAC 3596
QY 486 ----- 486
Db 3597 TTTTCTGTTAGTGTGATTACTCTCAATTTGTTCTTTCTTAAAAATCATATATCTGTG 3656
QY 487 -----Gln 487
Db 3657 TACTCCTCAAGAACTTGATTATATCTAAACGAGATTCTCATTTGGGAAAAATAAACACAG 3716
QY 488 ProGluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySer 507
Db 3717 CCAGAACACGATGAGAGGACCTGATGGAATCTAAATGTTGATCATGACGATGGGTCA 3776
QY 508 AspIleIleAlaAsnMetThrLysAlaProArgValLysTyrIleThrPheTyrGluAsp 527
Db 3777 GACATCATAGCTTAACATGACAAAGCACCAAGGTTAAGTACATACTTTTATGAAGAC 3836
QY 528 SerGluSerIleProGlyLysArgThrAlaValTyrGluLeuAspLysSerGlyTyr 546
Db 3837 TCTGAGAGCAATCCGGGGAAGAAACCGCAGTCTGGAGCTTGATAAAAAGTGGTAT 3893

RESULT 5
AAS01096
ID AAS01096 standard; cDNA; 328 BP.
XX
AC AAS01096;
XX
DT 31-MAY-2001 (first entry)
XX
XX Corn sterol acyltransferase LCAT EST sequence #3.
XX
XX lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
XX acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
XX nutritional supplement; dairy product; food product; salad dressing;
XX corn; Arabidopsis thaliana; expressed sequence tag; EST; ss.
XX
XX Zea mays.
XX OS
XX PN MO200116308-A2.
XX
XX 08-MAR-2001.
XX
XX 30-AUG-2000; 2000WO-US023863.
PF

```





Score: 345.00 Matches: 117  
Percent Similarity: 38.13% Conservative: 71  
Best Local Similarity: 23.73% Mismatches: 168  
Query Match: 11.64% Indels: 137  
DB: 12 Gaps: 17

US-09-651-651-5 (1-546) x ADF47831 (1-1872)

QY 44 ILEILEIPROGLYPHEALASERTHRGULEUARGALETTPSERILEULEAPCYV--- 62  
DB 445 ATTATGGTTCCTGGTGTATGACCTCAGATTAGAAAGTTGGTCGTTATATATGCTCG 504  
QY 63 ----ProlyThrProleu-----AapPheAsnProleuAspLeuValTrrpleu 77  
DB 505 ATTCCTACTTTAGGAACGCTCTTGGGAGCTGGCTGCTGCTGAGGCAATGCTTCTT 564  
QY 78 AapThrThrlyLeuLeuSerAlaValAsnCyETrrPheLySerCyMetValleuAapPro 97  
DB 565 GACAAAG-----CAATGCTGCTTGAAACATTAAATGCTTGATAA 603  
QY 98 TyrAenGlnThrAspHisProGluCyLySerArgProAspSerGlyLeuSerAlaIle 117  
DB 604 AAAACCGCTGGATCCGAAGGAATTAAAGCTGCGACAGCTCAGGGGTTTGAAAGCAGCT 663  
QY 118 ThrGluLeuAspProGlyTyrIleThrGlyProLeuSerThrValTrrpleuGluTrrpleu 137  
DB 664 GATTTTATATACGGGCTAT-----TGATTTGAGATTAAGTAAAT 705  
QY 138 LysTrpCyValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAapTrp 157  
DB 706 GAAACCTGCTGCATGCTGATGAGCCTAATACATGTTAAGTCTTCTTACGATG 765  
QY 158 ArgLeuSerProThrLyLeuGluGluArgAspLeuTyrPheHisLyLeuLeuLeuThr 177  
DB 766 CGGTTATCATATGCCAATTAGAGAAACGATTAATATTTTCAAAAGTTAAATAAGTTTC 825  
QY 178 PheGluThrAlaLeuLeuSerAlaArgGlyLyProSerIleValPheAlaHisSerMetGly 197  
DB 826 ATTGAGTACGCAACATTGTACTTAAGAAAGATGCTGTGATTTCTCATCTCATGGGT 885  
QY 198 AsnAsnValPheArgTyrPheLeuGluTrrpleuArgLeuGluIleAlaProLyHisTyr 217  
DB 886 TCACAGCTTACGACTATTTTAAAGTGGTGAAGCGGCTACGGAATGCTGA 945  
QY 218 LeuLyTrpLeuAspGlnHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuGly 237  
DB 946 CCGACTTGGGTTATGATCATATTGAAGCATTTTAATAATATCGGATCTTGAATGA 1005  
QY 238 SerValGluAlaIleLySerThrLeuSerGly-----ValThrPheGlyLeuProVal 255  
DB 1006 GCACCCAAACAGTGGCAGCGCTTTTATCGGTGAATGAATAACATACAGTATTTGTAAT 1065  
QY 256 SerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSer----- 270  
DB 1066 -----ACATTAAACATTTTGAATAAATTTTTCCTGAGAGAGCCATGATG 1116  
QY 271 -----LeuTrrpleuMetProPheSerLyAsnCyLySgLyAap 283  
DB 1117 GTTCGACTATGGAGAGATTAGTTCTTACTTCTTAAGAGAGCCATGTTGCTCCAGAT 1176  
QY 284 AsnThrSerTrpThrHisPheSerGlyGlyAlaAlaLySerAspLySerValTyrHis 303  
DB 1177 GATCTTAATCAAACTAATTTTCCATGTCATTAATTCGATATAGAGAAACATTTGAT 1236  
QY 304 CyAspGluGluGluTyrGln----- 310  
DB 1237 AAGACACACGATGATTTGACATAGATGATGATTAACATTTTAAATAATGTTACAGAT 1296  
QY 311 -----SerLyTrrSer-----GlyTrrProThrAsn 319  
DB 1297 GACGATTTTAAAGTCATGCTAGCGAAAAATTAATTTCCACGCTTGTGCTGAGCTGAAAA 1356  
QY 320 ILEILE-----AsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsn 337

DB 1357 GAAGTGTTAAAAAATTAACGAATATGCCGCTAAATGATNA----- 1395  
QY 338 MetThrSerMetGluCySgLyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAla 357  
DB 1396 --AATCCGCTAGAAACTAGTCTTCCCT----- 1419  
QY 358 ArgGlyThrLeuPheLyAlaIleGluAspTyrArgProAspSerLyAspMetLeuHis 377  
DB 1420 -----TATGCTCTGTATAGAAA----- 1437  
QY 378 GlnLeuLyLyLeuTyrHisAspAspProValPheAsnProLeuThrProTrrGluArg 397  
DB 1437 ----- 1437  
QY 398 ProProIleLyAsnValPheCyIleTyrGlyAlaHisLeuLyThrGluValGlyTyr 417  
DB 1438 -----ATTATTTGGCTTCACGGGGTCGGAACCAACATGAGAGGTTAT 1482  
QY 418 TyrPheAla-----ProSerGlyAspProTyrProAspAsnTrrpleIleThrAspIle 435  
DB 1483 TATTATTAATTAATCTGAGGGGCAACCTGCTGATGATTCCTCGGT----- 1530  
QY 436 IleTyrGluThrGluGlySerLeuValSerArgSerGlyThrValAlaAspGlyAsnAla 455  
DB 1531 -----AATGATGAAACAAAGTT--GAAATGATATGTTATGAT----- 1569  
QY 456 GlyProIleThrGlyAspGluThrValPro-TyrHisSerLeuSerTrrpCyLyAsnTr 475  
DB 1570 -----GATGGTGAATGGAACCTTACCAATATTAGCCCTTGGTGTGCAATAA 1620  
QY 475 PLeuGlyProLyValAsnIleThrMetAlaProGln 487  
DB 1621 GTTTGGCAACAAAGTTTAATCTGCTAATACAA 1657

RESULT 7  
ADT99409  
ID ADT99409 strand; cDNA; 1893 BP.  
AC ADT99409;  
XX  
XX 30-DEC-2004 (first entry)  
XX  
DE Fusarium lecitihin-cholesterol acyltransferase (LCAT) cDNA Seq 36.  
XX  
XX LCAT; lecithin-cholesterol acyltransferase; gene; bs;  
XX polyunsaturated fatty acid; PUFA; transgenic; acyltransferase;  
XX lipid metabolism; animal food; human food; cosmetic; pharmaceutical.  
XX  
OS Fusarium sp.  
XX  
XX  
FH Key Location/Qualifiers  
FT 1..1893  
FT /tag= a  
FT /product= "LCAT protein"  
XX  
XX  
XX MO2004087902-A2.  
XX  
XX 14-OCT-2004.  
XX  
XX 26-MAR-2004; 2004WO-BP003224.  
XX  
XX 31-MAR-2003; 2003DE-01014759.  
XX  
XX 17-OCT-2003; 2003DE-01048996.  
XX  
XX (UYBR-) UNIV BRISTOL.  
XX  
XX Renz A, Bauer J, Prentzen M, Soezer N, Keith S, Frazer T;  
XX Lazarus CM, Qi B, Abbadi A, Heinz E;  
XX  
XX WPI; 2004-737701/72.  
XX  
XX P-PSDB; ADT99410.  
XX

PT Preparing polyunsaturated fatty acids, useful in foods, animal feeds, cosmetics and pharmaceuticals, by growing organisms transformed with new PT sequences that encode specific acyltransferases.

PS Claim 1d; SEQ ID NO 36; 270bp; German.

XX This invention relates to a novel method for preparing polyunsaturated CC fatty acids (PUFAs) in a transgenic non-human organism by the CC introduction of any of 20 specific nucleic acid sequences, their CC derivatives or homologues thereof, followed by culturing and then CC harvesting the organisms. Specifically, it refers to nucleic acid CC sequences encoding polypeptides with acyltransferase activity, in CC particular, lysophosphatidic acid acyltransferase, glycerol-3-phosphate CC acyltransferase, diacylglycerol acyltransferase or lecithin-cholesterol CC acyltransferase. In addition, nucleic acids that encode proteins involved CC in lipid or fatty acid metabolism may be included such as acyl-CoA CC dehydrogenase, fatty acid hydroxylase or acyl-acyl carrier protein CC desaturase. The present invention describes long chain multiply CC unsaturated fatty acids (carbon 18, C20, C22 or C24 acids) that are CC recovered as oils, lipids or free fatty acids including arachidonic, CC eicosapentaenoic and docosahexaenoic acids, where the organisms are CC microorganisms, non-human animals or plants. Furthermore, these fatty CC acids, oils or lipids, including the derived compositions, can be used in CC animal or human foods, cosmetics and pharmaceuticals, as well as for CC synthetic intermediates. This polynucleotide sequence is the Fusarium CC gramineum lecithin-cholesterol acyltransferase (LCAT) cDNA clone (coding CC sequence) of the invention.

XX Sequence 1893 BP; 477 A; 473 C; 509 G; 434 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1,68e-23	Length:	1893
Score:	339.00	Matches:	146
Percent Similarity:	38.23%	Conservative:	78
Best Local Similarity:	24.91%	Mismatches:	224
Query Match:	11.43%	Indels:	138
DB:	13	Gaps:	27

US-09-651-651-5 (1-546) x ADT99409 (1-1893)

QY 20 LeuilecysgllygylatgthrAlaValGluaspGluThrGluPheHis---GlyaspTyr 38  
 DB 319 CTGTTGAGGCGCGGAGCATTCGCCAATGTAGACAGCCATCTCTGTTGGGAAG 378  
 QY 39 SerlyleuSerGly-----llelleleProGlyPheAlaSer 51  
 DB 379 GCTGATCCGAGGCTTAGAGATTCACCATCTATGATCATGATACCTGGGTATCTCA 438  
 QY 52 ThcGlnleuArgAlaTTPSerle-----LeuaspCysProTyrThrProleuaspPhe 69  
 DB 439 ACTGAGCTCGAATGTGGCGGTACGGCTAATATCTCGAAACCTTCTT----- 486  
 QY 70 AsnProleuaspLeuValTTPLeuaspThrThryleuLeuSerAlaVal----- 86  
 DB 487 -----AGAAACGACCTTGGGTAGTTGACCAATGATGAGAGCTCTGGTATGACAG 540  
 QY 87 AsnCytrPheleuValMetValleuaspProTyrAsnGlnThrAspHisProGluCys 106  
 DB 541 GAGGTTTGAAAGAACACATGATCTGCACAAAGAGAGCGGCTTGACCCGCTGACGTA 600  
 QY 107 LyseSerArgProaspSerGlyleuSerAlaIleThrGluLeuaspProGlyTyrIleThr 126  
 DB 601 AAGTTGAGGCTGCCCAAGGGTTCGATGCGACCGATTCTTCATCAGCGGATAT----- 654  
 QY 127 GlyProleuSerThrValTTPLeuaspThrProleuaspTyrCysValGluPheGlyIleGlu 146  
 DB 655 -----TGATCTGAGCGCAAAATCTTTAGAAATCTCGCATCCATCGCGTCAAC 702  
 QY 147 AlaAsnAlaIleValAlaValProTyrAspTTPArgLeuSerProThrIleLeuGluGlu 166  
 DB 703 CCAAGCAACTGCTTACAGGCTGCTTACGATTCGGCTTGTGATCCCAACCTTGAGGTA 762  
 QY 167 ArgAspLeuTyrPheHisIleLeuLeuValLeuThrPheGlnThrAlaLeuValLeuArgGly 186

DB 763 CGGAGCCGCTACTTCTGCTGGCTAAAGTCCATATGAAATCGCGTGGCCACTGAGAC 822  
 QY 187 GlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeuGlu 206  
 DB 823 AAAAAGTGTCTCTCCATCAGACAGTATGGGAGCAAGTCTTACTATTTTCTCCAC 882  
 QY 207 TTPLeuArgLeuGluIleAlaProlyshIstYrLeuIleTTPLeuaspGlnHisIleHis 226  
 DB 883 TGGGTCTACATCGAAGAGCGGACCGGTGGCCGATTTGGATTGGATGCTCATTTGAC 942  
 QY 227 AlaTyrPheAlaValGluIleAlaProleuLeuGlySerValGluAlaIleLyseSerThrLeu 246  
 DB 943 GCGTGAATCAACATCAGCGGATGCAATGCTTGAAGCATGAAGATTGACCGCTGCTC 1002  
 QY 247 SerGly-----ValThrPheGlyLeuProVal 255  
 DB 1003 TCCGCGAGATGGCGACACAGCTCAACTGAAACCGTTCGCTATTACGCGCTG----- 1056  
 QY 256 SerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerIleTyr----- 272  
 DB 1057 -----GAAAGTCTTGAGTAAAGAGAGAGAGCGAGATCTTTCCGCGCATG 1104  
 QY 273 -----LeuMetProPheSerIleAsnCylyaspGlyAspAsnThrSerTyr 287  
 DB 1105 CCCGGATATCTCTCATGTGGCCCATCGCGGCACTCTGTATGGGTTACTTGACCTGG 1164  
 QY 288 ThrHisPheSerGlyGlyAlaIleAlaIleValAspIleValIstCysaspGluGlu 307  
 DB 1165 GCTCCA-----GACGACTTG 1179  
 QY 308 GluTyrGlnSerIstYrSerGlyTyrProThrAsnIleleAsnIleGluIleProSer 327  
 DB 1180 CCAGCGCCAGAACCGTTCAATATGGA-----TCTCTCTGAACTTTAGAGTCGCTTCG 1230  
 QY 328 ThrSerValThrGluThrIleAlaValAsnMetThrSerMetGluCysGluLeuProThr 347  
 DB 1231 AACTGACAACTCTCGATGCT-----AACTTTTACC--GTGAGAAAGGTGTCTTAT 1281  
 QY 348 LeuLeuSerPheThrAlaArgIleLeuAlaAspGlyThrLeuPheIleAlaIleGluAsp 367  
 DB 1282 TTGCTTAAACAAAG-----GAGGAC 1302  
 QY 368 TyrAspProAsp-----SerlyAspArgMetLeuHisGlnLeuIstYs 381  
 DB 1303 TGGTATCAAGACCAAGATCAAGGCGAGTATTTCTCGGGCAATTCCTCATTCACATAGATGAG 1362  
 QY 382 Leu-----TyrHisAspAspProVal-----PheAsnProleu---ThrProTyrGlu 396  
 DB 1363 GTCGAAGCCAAATGAGATGACCCCAAGAGAGTGAATCATCTCTCGAGAGCGCATTCGCA 1422  
 QY 397 ArgProProIleuValAsnValPheCysIleTyrGlyAlaHisIleuValThrGluValGly 416  
 DB 1423 CTGCTCTTGGCTCGCAAGATCTACTGCTTTATAGTCTTGGAAGAACCGACGAGCGAGG 1482  
 QY 417 TyrTyrPheAlaProSerGlyIstYrProTyrProAspAsnTTPIleIleThrAspIleIle 436  
 DB 1483 TACTTATTAAGCCACCGATCAGCATCATTTGACCAACCTCAACATCACAATATGATAGAC 1542  
 QY 437 TyrGluThrGluGlySerIleValSerArgSerGlyThrValValAspGlyValAsnAlaGly 456  
 DB 1543 GCGTATTCGCAAGAGAC-----GTGATCAATGCGCGTTGTC 1578  
 QY 457 ProIleThrGlyAspArgIleThrValProTyrHisSerLeuSerTyrCysIstYsAsn----- 474  
 DB 1579 ATGGGCGAGGAGATGGTACCGTAACTCTCTGATCAGAGGATCATGTGATATCATGCG 1638  
 QY 475 Tyr-----LeuGlyProIstYsValAsnIleThrMetAlaPro--- 486  
 DB 1639 TGGAAATGAAACGCTCAACCCAGGAGCGGTCAAGGTTACAGTTGTGAGATGCGCTCAC 1698  
 QY 487 GluProIleHis-----AspArgIstAspVal----- 495

DB 1699 GAGCGGACCGCTTCATCTCGAGGAGGGCCGCGACGGCCGACACGTTGACATCTTG 1758  
QY 496 -----HisValGluLeuAsnValAspHisGluHisGlySerAspIle 509  
DB 1759 GGGCGATACACCTGGAACGAGTTGCTTACGATGCGACGCGCAAGGTGACACGATT 1818  
QY 510 -----IleAlaAsnMetThrLys---AlaProArgValIleTyrIleThrPhe 524  
DB 1819 ACGAAGTATGTTGTGTGACCAATCAAAAGATATGATCCAGGGTTAAG-----ATT 1869  
QY 525 TyrGluAspSerGluSer 530  
DB 1870 TACGATGATGACGAGACT 1887  
RESULT 8  
ADM20631  
ID ADM20631 standard; DNA; 2326 BP.  
XX  
AC ADM20631;  
XX  
DT 07-APR-2005 (first entry)  
XX  
DE Y. lipolytica phospholipid:diacylglycerol acyltransferase (PDAT) DNA.  
XX  
KM malnutrition; eating-disorders-gen.; nutritional disorder; inflammation;  
XX antiinflammatory; hypercholesterolemia; antilipemic; metabolic disorder;  
XX oil; food; phospholipid:diacylglycerol acyltransferase; PDAT; ds; gene.  
XX  
OS Yarrowia lipolytica.  
XX  
FH Key Location/Qualifiers  
FT CDS 274..2220  
FT /tag= a  
FT /product= "Y. lipolytica phospholipid:diacylglycerol  
FT acyltransferase (PDAT) protein"  
FT /EC\_number= "EC 2.3.1.158"  
XX  
PN MO2005003322-A2.  
XX  
PD 13-JUN-2005.  
XX  
PF 02-JUL-2004; 2004MO-US021932.  
XX  
PR 02-JUL-2003; 2003US-0484599P.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.  
XX  
PI Yadav NS, Zhang H;  
XX  
DR WPI: 2005-091802/10.  
XX P-PSDB; ADM20632.  
XX  
PT New nucleic acids encoding diacylglycerol acyltransferase and  
PT phospholipid:diacylglycerol acyltransferases useful for altering  
PT polyunsaturated fatty acid content in triacylglycerols in oleaginous  
PT yeasts.  
XX  
XX Example 3; SEQ ID NO 45; 127bp; English.  
XX  
CC The invention relates to novel isolated nucleic acids encoding  
CC diacylglycerol acyltransferases (diacylglycerol O-acyltransferase, DGAT,  
CC EC 2.3.1.20) and phospholipid:diacylglycerol acyltransferases (PDAT, EC  
CC 2.3.1.158) and the encoded polypeptides. The nucleic acids and methods of  
CC the invention may be useful for increasing triacylglycerol (TAG) content  
CC in transformed host cells, especially in oleaginous yeasts. The  
CC triacylglycerol produced preferably comprises long-chain omega-3/omega-6  
CC polyunsaturated fatty acids (PUFAs). The methods allow production of  
CC PUFAs with improved nutritional qualities which may be utilized as  
CC dietary supplements, particularly for infant formulas, patients  
CC undergoing intravenous feeding or for preventing or treating  
CC malnutrition, inflammation or high cholesterol. The purified PUFAs may  
CC also be incorporated into cooking oils, fats, margarine or other food  
CC products. Certain PUFAs are recognized as essential fatty acids that

CC cannot be synthesized de novo in mammals and must be supplied as part of  
CC the diet or derived from linoleic acid or alpha-linolenic acid. There are  
CC a variety of commercial sources of PUFAs from natural sources. However,  
CC there are disadvantages associated with these methods of production e.g.  
CC heterogenous oil compositions, pollution, availability fluctuation due  
CC to the weather and expense. The current sequence is that of the Yarrowia  
CC lipolytica phospholipid:diacylglycerol acyltransferase (PDAT) DNA of the  
CC invention.  
XX  
SQ Sequence 2326 BP, 560 A, 593 C, 629 G, 543 T, 0 U, 1 Other;

Alignment Scores:  
Pred. No.: 5,186-22 Length: 2326  
Score: 325.50 Matches: 119  
Percent Similarity: 38.14% Conservative: 66  
Best Local Similarity: 24.54% Mismatches: 202  
Query Match: 10.98% Indels: 99  
DB: Gaps: 13

US-09-651-651-5 (1-546) x ADM20631 (1-2326)

QY 27 AlaValGluAspGluThrGluPheHisGlyAspTyrSerLysLeuSerGlyIleIleIle 46  
DB 709 GCTGTGGCGACAGCTATGAAAGCGAGGACTCAACGCCAAGTACCAGGTGCTGGTG 768  
QY 47 ProGlyPheAlaSerThrGlnLeuArgAlaTrpSerIle-----LeuAspCyPro 63  
DB 769 CCCGGCGTCATCTCCACGGGACTGGAGAGCTGCTCCCTCGAGAGGAACCGAGAGTCTCC 828  
QY 64 TyrThrProLeuAspPheAsnProLeuAspLeuValTyrLeuAspThrThrLysLeuLeu 83  
DB 829 ACCGAGTCGCACTTCAGAAAGCAATGTGGGCTCTGTGCATGATCCGATGCTGCTG 888  
QY 84 SerAlaValAsnCyThrPheLysCysMetValLeuAspProTyrAsnGlnThrAspHis 103  
DB 889 CTGGAACAAGTACTGCTGCTGCGAAGACCTGATGCTGACACAGACAGACCGGTAAACCT 948  
QY 104 ProGluCyLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGly 123  
DB 949 CCCCAATTCACACTCGAGCCGCCCAAGGATTGGCTCGCGGACTTCTTTATGGCAGGC 1008  
QY 124 TyrIleThrGlyProLeuSerThrValTyrLysGluTyrPheLysTyrPheValGluPhe 143  
DB 1009 TAC-----TGGCTGTGGAACAAGCTGCTGGAACCTGGCTGTATT 1050  
QY 144 GlyIleGluAlaAsnAlaIleValAlaValProTyrAspThrArgLeuSerProThrLys 163  
DB 1051 GGATACGATACGATACATATGCTCTGCGGCGTACGACTGGAGACTGCTTACCCCTGAT 1110  
QY 164 LeuGluGluArgAspLeuTyrPheHisLysLysLeuLysLeuThrPheGluThrAlaLeuLys 183  
DB 1111 TTGGAGCACCGGACGAGTACTTCTCCAAAGCTCAAGCTCAATGAGAGACTAAGCT 1170  
QY 184 LeuArgGlyLysProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArgTyr 203  
DB 1171 ATGACAGGTGAAGAAACAGCTTCTGACGGGCGCATTCATGAGGCTCCAGGTCACTTCTTAC 1230  
QY 204 PheLeuGluTyrPheLysArgLeuGluIleAlaProLysHisTyrLeuLysThrLeuAspGln 223  
DB 1231 TTCATGAATGGGCTGAGGCGGAGGATGTGAGAGAGAGGTCCTCAACTGCTCAATGAC 1290  
QY 224 HisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuGlySerValGluAlaIleLys 243  
DB 1291 CATATTGAATCTCTTGTGCACTTTCCGGCTCGATGCTGGGATACCCCAAGACCTGGT 1350  
QY 244 SerThrLeuSerGly-----ValThrPheGlyLeuProValSerGluGly 258  
DB 1351 GCTCTTCTGTGGGAATGAAGATACCGGACGCTGAACGCGATGGCTGTATGGA 1410  
QY 259 ThrAlaArgLeuLeuSerAsnSerPheAlaSerLeu-----TyrLeuMetPro 275  
DB 1411 CTGAGACAGTCTTCTCTGACGAGAGCGAGCGATCTGCTCGAACAATGGGAGGAATT 1470

```

Oy 276 PheSerIysAsnCyIysGlyAspAsnThrSerTrpThrHisPheSerGlyIValAla 295
    |||
    |||
    |||
Db 1471 GCTTCATGATCTCCCAAGGGTGAAGGCTATCGGGGTGATCTTGAGGCCCTC --- 1527
Oy 296 LysLeuAspLysArgValYrHisCysAspGluGlu --- 307
    |||
    |||
    |||
Db 1528 -----GATGACGAGCCCGCCGCAATGTCACCTTTGGC 1560
Oy 308 -----GluTyrGlnSerLys-TyrSerGlyTrp-- 316
    |||
    |||
    |||
Db 1561 AACTTCATCAAGTTCAAGAGCTCTTACCGAGTACTCTGTAAGAACCTCACATGAT 1620
Oy 317 ----ProThrAsnIleIleAsnIleGluIleProSerThrSerValThrGluThrAla 335
    |||
    |||
    |||
Db 1621 GAACCCGTGACTTCTGTATTTCTCACTCTCCGAGGTGTTTGAAACCAACCGAGGCT 1680
Oy 335 uValAsnMetThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArg 355
    |||
    |||
    |||
Db 1681 G-----CTTACTCTCTTTGAA-----TTGCCAAGACTGAAAGCAGGTTGAGCAGAT 1728
Oy 355 uLeuAlaAsp-----GlyThrLeuPheLysAlaIleGluAspTyrAspProAs 371
    |||
    |||
    |||
Db 1729 GAGAAGGACCTTCTACTCGAGCAACCTCTGGAAG----- 1765
Oy 371 pSerLysArgMetLeuHisGlnLeuLysLysLeuTyrHisAspAspProValPheAsnPr 391
    |||
    |||
    |||
Db 1766 -----CTGCTCTCTCCC 1776
Oy 391 oLeuThrProTrpGluArgProProIleLysAsnValPheCysIleTyrGlyAlaHisLe 411
    |||
    |||
    |||
Db 1777 AATG-----CCCCGATCTC-AAGATCTACTGCTTCTTATGAGTGGGTAA 1820
Oy 411 uLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAsnTrpI 431
    |||
    |||
    |||
Db 1821 GGATACCGAGCGACCTACTACTCCAGAGATGACCCCAATCCAGAGCAACCACTTGA 1880
Oy 431 eIleThrAspIleIleTyrGluThrGlySerLeuLysValSerArgSerGlyThrVal 451
    |||
    |||
    |||
Db 1881 CGTCAGATCGCTCGAAGACGACCTGATGGTGTCTTATGGGTGAG----- 1926
Oy 451 lAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSerLeuSerTr 471
    |||
    |||
    |||
Db 1927 -----GGCGATGGAACCGTCTCCCTTGATCCCATTCAT 1961
Oy 471 pCyLeuAsnTrp 475
    |||
    |||
    |||
Db 1962 GTGTCACCGATCG 1974

```

```

PR 29-MAY-2002; 2002US-0383889P.
PR 20-JAN-2003; 2003SE-00000142.
PA (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.
PI Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;
DR WPI, 2004-053268/05.
DR P-PSDB; ADF47835.
XX
XX
XX New nucleotide sequence encoding an improved acyltransferase polypeptide,
XX useful for producing structured lipids or fat-soluble molecules, in
XX removing undesirable fat or in modifying lipids in animal or plant raw
XX material.
XX
XX Claim 5; SEQ ID NO 20; 91pp; English.
XX
XX The present invention describes a nucleotide sequence (I) derived from a
XX nucleotide sequence encoding an acyltransferase polypeptide comprising at
XX least one membrane-spanning region. (I) encodes an improved active
XX membrane independent acyltransferase polypeptide in which at least one
XX amino acid residue of the membrane-spanning region has been deleted
XX and/or substituted as compared to the original acyltransferase
XX polypeptide, where the encoded active membrane independent
XX acyltransferase polypeptide can produce fatty acid esters and/or fatty
XX acid thioesters such as triacylglycerols, diacylglycerols, waxes,
XX monacylglycerols, phospholipids, glycolipids, lysolipids, waxes,
XX acylated carbohydrates and acylated amino acids. Also described: (1) a
XX nucleotide sequence molecule (II) comprising at least one promoter region
XX which functions in a host, where the promoter region is operably linked
XX to at least one (I), which is operably linked to at least one non-
XX translated region which functions in a host; (2) a vector comprising (II)
XX ; (3) a host cell comprising (II) or the vector; (4) producing an
XX improved active membrane independent acyltransferase polypeptide,
XX comprising providing the above host cell and a growth medium preparing a
XX host cell culture, culturing the host cell culture and harvesting the
XX host cell culture and recovering the improved active membrane independent
XX acyltransferase polypeptide; (5) a polypeptide obtained by the method in
XX (4), where the polypeptide at least is an improved active membrane
XX independent acyltransferase polypeptide; (6) an oligonucleotide
XX specifically hybridising to (I) under stringent conditions; and (7) a kit
XX comprising the above polypeptide and a stabiliser. The nucleic acid
XX molecule (I) and the polypeptide are useful in producing structured
XX lipids or fat-soluble molecules, in removing undesirable fat or in
XX modifying lipids present in animal and plant raw material. The present
XX sequence is used in the exemplification of the present invention.
XX
XX Sequence 2004 BP; 567 A; 373 C; 544 G; 520 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.36e-19 Length: 2004
XX Score: 300.50 Matches: 134
XX Percent Similarity: 33.22% Conservative: 63
XX Best Local Similarity: 22.60% Mismatches: 239
XX Query Match: 10.13% Indels: 157
XX DB: 12 Gaps: 23
XX
XX US-09-651-651-5 (1-546) x ADF47835 (1-2004)
Oy 44 lIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSerIleLeuAspCysPro 63
    |||
    |||
    |||
Db 346 GTTTAGTCCCGGAATCGTACCGGCGGCTTGAGCTTTGGGAAGCCCACTTGTCT 405
Oy 64 TyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThr--ThrLysLeu 82
    |||
    |||
    |||
Db 406 CATGCTCTTTT-----CGTAAACGCTTGCGGTGTGACTTTCAGAAATC 453
Oy 83 LeuSerAlaValAsnCyTrpPheLysCysMetValLeuAspProTyrAsnGlnThrAsp 102
    |||
    |||
    |||
Db 454 TTAGAGAGCCCATTTGTGGTGGACACTTGTGCATGAGCAATGAACGAGGCTGCAC 513
Oy 103 HisProGluCysLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeuAspPro 122
    |||
    |||
    |||

```

```

Db      514 CCGCGGGGATTCGGGTCAGACCAAGTCCGGGCTTGTGACGTATTTCGCTCT 573
Qy      123 G|Y|T|T|e|T|h|e|G|L|P|r|o|e|u|s|e|r|T|h|r|V|a|I|T|P|L|y|S|G|U|T|P|L|e|u|L|y|T|P|C|y|S|V|a|G|U| 142
Db      574 GGGTATTTTC-----GTTTGGCGGGTTCTTATCGAGATCTGGGTAA 615
Qy      143 P|h|e|g|l|l|e|g|l|a|a|n|l|a|e|n|l|a|l|e|v|a|l|a|P|r|o|T|A|S|P|T|P|A|r|g|h|e|u|s|e|r|P|r|o|h|r 162
Db      616 ATCGGCTACGAAGGGAAGAACATGTATATGCGCTTACGACTCGAGCGCTTTCGTTTCA 675
Qy      163 L|y|L|e|u|L|G|U|L|A|S|P|L|e|u|T|P|h|e|n|l|S|y|L|e|u|L|y|L|e|u|T|h|r|P|h|e|G|U|T|h|r|a|l|e|u| 182
Db      676 AACACGGAAGTAAGGATCAGACACTTACGACATTAATAATCAATCGACGAATATGAC 735
Qy      183 L|y|L|e|u|A|T|G|L|y|L|P|r|o|S|e|r|l|e|v|a|l|P|h|e|a|l|a|---H|l|S|e|r|T|e|G|L|y|A|n|S|e|n|V|a|l|P|h|e 201
Db      736 ATACCAACGCTAACAAAGAAAGTAGTGCTGTCACATTCATTAATGGGGTAAATCTAATTC 795
Qy      202 A|T|G|T|P|h|e|L|e|U|T|P|h|e|U|A|T|G|L|U|L|e|-----A|A|P|P|L|y|S 215
Db      796 CTACACTTCTCAATGGGTGGAAGCCCTGTTCCGATGGCGGCGCGCGCTCCA---- 852
Qy      216 H|S|T|Y|L|e|u|L|y|T|P|h|e|u|A|P|G|L|N|h|S|I|e|N|h|S|A|L|e|T|Y|P|h|e|A|V|a|G|L|U|L|A|P|r|o|L|e|u| 235
Db      853 -----GGGTGGTCGACAAAGCATATCAAGCTATCATGAAACATCGAGCGCGCTTT 903
Qy      236 L|e|U|L|y|S|e|r|V|a|G|U|L|a|l|e|L|y|S|e|r|T|h|r|L|e|u|S|e|r|L|y|-----V|a|l|T|h|r|P|h|e 251
Db      904 CTCGGTGTTCGAAAGCGTTAGCGGGATGCTGTCGGCGAAGCAAGACGTTGCTTTT 963
Qy      252 -----G|L|y|e|u|P|r|o|V|a|L|S|e|r|G|U| 257
Db      964 TTCAGAGCCATGGCGCTGTTGATAGATTCAAGAAATTCAGAGCTCCAAACTCAGAA 1023
Qy      258 G|Y|T|h|r|A|A|T|G|L|e|u|L|e|u|S|e|r|S|e|n|S|e|r|P|h|e|A|S|e|r|S|e|r|U|T|P|h|e|u|T|P|h|e|u|T|P|h|e|S|e|r 277
Db      1024 CACATGATCGCT---GTAGTGCAGACTTGGGATTCGTATCTCATTTGCTTCTCT----- 1074
Qy      278 L|y|S|e|r|C|y|L|y|S|e|r|P|h|e|n|T|h|r|S|e|r|T|P|h|S|e|r|G|L|y|----- 292
Db      1075 -----AAAGCGGAGACACCATTTGGGGGATTTGATTCTCCCTGAAAGATTTC 1125
Qy      293 -----G|Y|A|L|a|L|y|L|y|S|e|r|P|L|y|A|T|G|V|a|l|T|h|r|H|S|e|r|S|e|r|G|U|L|U| 307
Db      1126 GAAACGAAACCATCGTAAATTAACACACGAAACCTCAGTTCGTATGAACAGAA 1185
Qy      308 G|U|T|Y|G|L|S|e|r|L|y|S|e|r|G|L|Y|T|P|-----P|r|o|h|r|A|n 319
Db      1186 ACGAAATACGGAAGATTAATCATTCGGAAGACAGCTTCGAGCAACATTCCTCAAAAT 1245
Qy      320 I|l|e|l|a|S|e|r|l|e|G|U|L|e|P|r|o|S|e|r|T|h|r|S|e|r|V|a|l|T|h|r|G|U|T|h|r|a|l|e|u|V|a|l|a|n|S|e|r|T|h|r 339
Db      1246 CTCACAAATCAGCATCTTCAAAAAGATTAATTATTCGAC-----ACC 1287
Qy      340 S|e|r|T|e|G|L|Y|S|G|L|Y|----- 344
Db      1288 GGTGCAAAATTCGCGTGAAGATTGGAGCGAGTACGCTAAATTTAGCAAGAAACATTATA 1347
Qy      345 ---L|e|u|P|r|o|h|r|L|e|u|L|e|u|S|e|r|P|h|e|T|h|r|A|A|T|G|L|U|L|e|u|L|a|S|e|r|G|U|L|y|T|h|r|L|e|u|P|h|e|L|y|S 363
Db      1348 AAACCTTGCTGATTAACAAACGCTACACAGCTGAACGTTATCGATCTTTTACGCTTTGG 1407
Qy      364 A|A|l|l|e|G|U|A|S|P|L|y|T|P|h|e|P|r|o|A|S|e|r|L|y|A|T|G|L|U|L|e|u|H|S|e|r|l|e|u|L|y|S|e|r|L|y|T| 383
Db      1408 GCA-----CCAAAGACGATGAACGGGCGGAGCCCATTTCTCA 1446
Qy      384 H|S|-----A|S|A|P|P|P|r|o|V|a|l|P|h|e|A|n|P|r|o|L|e|u|h|r|P|r|o|T|P|G|L|U|A|T|G 397
Db      1447 CATGGAATAGCGGATGATCTTGTGATGACCCAAATATATATATTAACAATTTGTGGAAT 1506
Qy      398 P|r|o|-----P|r|o|L|y|S|e|r|V|a|l|P|h|e|C|y|L|e|r|G|L|Y|A| 409
Db      1507 CCGTTAGAGACTAAACTACAGATGCTCCAAACATGAGATTTTCTTTGTAAGGGGTT 1566

```

```

Qy      410 H|S|L|e|u|L|y|T|h|r|G|U|L|V|a|G|L|Y|T|P|h|e|A|L|P|r|o|S|e|r|G|L|Y|S|P|r|o|T|Y|P|r|o|A|S|e|r|P|h|e 429
Db      1567 GGACTCTTACGAAAGATCGTATGTATATAGACTGTACAAATC----- 1611
Qy      430 T|P|h|e|l|l|e|r|h|A|S|P|l|e|T|G|U|T|h|r|G|U|L|y|S|e|r|V|a|S|e|r|A|T|S|e|r|G|L|Y|T|h|r 449
Db      1612 -----G|A|T|A|A|T|G|C|A|C|A|G|A|T|A|T|T|C|G|T|T|A|G|A|A|T|T|G|A|C|A|G 1647
Qy      450 V|a|l|A|S|e|r|G|L|Y|A|n|A|G|L|Y|P|r|o|-----I|l|e|h|r|G|L| 460
Db      1648 TCAGCGGAGGGGAACGCTGTCGGGAGCTTACGTGCGGCTTTATTTTGTGATGCT 1707
Qy      461 A|P|G|U|T|h|r|V|a|l|P|r|o|T|h|r|H|S|e|r|L|e|u|S|e|r|T|P|-----C|y|L|y|S|e|r|T|P|h|e|U|L|Y|--- 477
Db      1708 CATGGAAGTATGACCGGTTTGTAGTCCGGGCTTATATGTGTCGAAAGGTCGAAAGGAAA 1767
Qy      478 -----P|r|o|L|y|V|a|L|a|n|l|e|T|h|r|M|e|t|A|l|P|r|o|G|l|n|P|r|o|G|l|n|H|S|e|r|G|L|Y|--- 492
Db      1768 ACTAGTTTAAATCCGCTGCGAGTGAGACGTATATACGAGGTATAAACATTAACACCG 1827
Qy      493 -----S|e|r|A|P|V|a|l|H|S|e|r|V|a|G|L|U|L|e|u|A|n|A|S|e|r 502
Db      1828 GGAAGTTGCTTGAGGAAAGGCTTTGAAAGTGGGCTCATGTT----- 1872
Qy      503 H|S|G|U|H|S|G|L|S|e|r|A|P|l|e|l|e|A|A|n|S|e|r|T|h|r|L|y|S|A|L|P|r|o|A|T|G|V|a|l|L|y|T|h|r|l|e 522
Db      1873 -----G|A|T|A|T|T|G|G|G|A|A|T|G|C|A|T|T|G|A|G|A|T|T|T|G|A|G|G|G|G| 1917
Qy      523 T|h|r|P|h|e|T|Y|G|L|U|A|S|e|r|S|e|r|L|e|P|r|o|G|L|Y|A|T|G 535
Db      1918 GCGGCCGAGCTTCTGGGCTGGAGATGCGCGGTATCGG 1956

RESULT 10
ADF47817
ID      ADF47817 standard; DNA; 1701 BP.
XX
AC      ADF47817;
XX
DT      26-FEB-2004 (first entry)
XX
DE      Yeast membrane independent acyltransferase DNA sequence SEQ ID NO:2.
XX
KW      acyltransferase; enzyme; membrane-spanning region;
XX      active membrane independent acyltransferase; fatty acid ester;
XX      fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;
XX      phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;
XX      acylated amino acid; structured lipid; fat-soluble molecule; yeast; gene;
XX      ds.
XX
OS      Saccharomycetes cerevisiae.
XX
PN      WO2003100044-A1.
XX
PD      04-DEC-2003.
XX
PF      28-MAY-2003; 2003WO-SE000870.
XX
PR      29-MAY-2002; 2002SE-00001581.
XX      29-MAY-2002; 2002US-0383889P.
XX      20-JAN-2003; 2003SE-00000142.
XX
PA      (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.
XX
PI      Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;
XX
DR      MPI; 2004-053268/05.
XX
PT      New nucleotide sequence encoding an improved acyltransferase polypeptide,
XX      useful for producing structured lipids or fat-soluble molecules, in
XX      removing undesirable fat or in modifying lipids in animal or plant raw
XX      material.
XX

```







DB 1552 -----GAGCTCCCAACATGAAA---ATCTATTGTATATACCGGCTGAACAACCCAACT 1602  
QY 414 G|U|V|A|G|Y|T|Y|T|P|H|A|P|P|S|E|R|G|Y|L|Y|P|P|T|Y|R|P|H|A|S|E|R|T|H|I|E|T|H| 433  
DB 1603 GAAAGGCGCATAT----- 1614  
QY 434 A|S|P|I|E|L|E|T|Y|R|G|L|U|R|G|I|U|G|Y|S|E|R|L|E|U|A|S|E|R|A|R|S|E|R|G|Y|T|H|V|A|L|A|A|P|G|I|Y 453  
DB 1615 -----G|T|A|T|T|A|A|G|A|A|G|A|G|A|T|G|A|C|T|C|T|G|C|T|G|T|G|A|T|T|G|A|C|A|T|C|G|A|T|C|G|A 1668  
QY 454 A|S|N|A|L|G|I|P|R|O|L|E|-----T|H|G|Y|A|S|P|L|U|T|H|V|A|L|P|R|O|T|Y|R|H|S|E|R|L|E|U|S|E|R 470  
DB 1669 A|G|C|A|G|C|A|A|C|C|T|G|T|A|T|T|C|T|C|A|C|C|G|G|G|A|C|G|A|A|C|C|T|T|C|G|T|C|G|G|C|G|C|A|T|T|C 1728  
QY 471 T|P|C|Y|L|Y|S|A|S|N|T|P|-----L|E|U|G|I|P|R|O|L|Y|S|-----V|A|L|A|S|I|E|T|H| 483  
DB 1729 A|T|G|T|C|A|C|A|A|A|G|G|G|C|C|A|G|G|G|T|G|T|T|C|A|C|C|G|T|A|C|A|C|C|T|G|C|G|G|A|T|T|A|C|G|T|T|A|C|T 1788  
QY 484 -----M|E|T|A|P|R|O|G|I|N|P|R|O|G|I|U 489  
DB 1789 A|T|G|T|G|A|A|T|G|A|A|A|C|A|C|C|A|G|C|A|G|A|T 1815  
RESULT 12  
AAC64440  
ID AAC64440 standard; DNA; 1986 BP.  
XX AAC64440;  
AC AAC64440;  
XX  
XX 12-FEB-2001 (first entry)  
DT  
XX  
DE Saccharomyces cerevisiae PDAT ORF nucleotide sequence SEQ ID NO:4a.  
XX  
KM PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;  
KW EST; expressed sequence tag; fatty acid; oil content; de.  
XX  
XX Saccharomyces cerevisiae.  
OS  
XX  
PN W0200060095-A2.  
PD 12-OCT-2000.  
XX  
PF 28-MAR-2000; 2000WO-EP002701.  
XX  
PR 01-APR-1999; 99EP-00106656.  
PR 10-JUN-1999; 99EP-00111321.  
PR 07-FEB-2000; 2000US-0180687P.  
XX  
XX (BADI ) BASF PLANT SCI GMBH.  
PA  
XX  
PI Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Stymne S;  
XX WPI, 2000-665012/64.  
DR P-PSDB; AAB24265.  
XX  
XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic  
PT pathway for triacylglycerol production and DNS encoding them, useful for  
PT producing triacylglycerol, or for transforming any cell or organism to  
PT increase oil content.  
XX  
XX  
PS Claim 6; Page 71-74; 97pp; English.  
XX  
XX The present invention describes an enzyme for catalysing (in an acyl-CoA-  
CC independent reaction) the transfer of fatty acids from phospholipids to  
CC diacylglycerol in the biosynthetic pathway for the production of  
CC triacylglycerol (TAG). The enzyme is designated as  
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the  
CC nucleotides encoding them are useful for producing TAG and/or TAG with  
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for  
CC transforming any cell or organism in order to be expressed in this cell  
CC or organism and result in an altered, preferably increased oil content of  
CC this cell or organism. The present sequence represents the yeast  
CC (Saccharomyces cerevisiae) PDAT ORF (open reading frame) nucleotide

CC sequence  
XX  
SQ Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1,7e-19 Length: 1986  
Score: 299.50 Matches: 119  
Percent Similarity: 38.51% Conservative: 77  
Best Local Similarity: 23.38% Mismatches: 188  
Query Match: 10.10% Indels: 125  
DB: 3 Gaps: 24  
US-09-651-651-5 (1-546) x AAC64440 (1-1986)  
QY 37 A|S|P|Y|T|-----S|E|R|L|E|U|S|E|R|G|Y|L|I|E|L|I|E|P|R|O|G|I|P|H|A|S|E|R|T|H|I|N| 53  
DB 496 G|A|T|T|A|T|A|T|G|A|G|G|C|A|A|C|A|T|C|G|T|G|T|A|A|T|G|T|C|G|T|C|A|T|T|C|T|A|C|G|G|A 555  
QY 54 L|E|U|A|G|A|L|A|T|P|S|E|R|I|L|E|U|A|P|R|O|T|Y|R|T|H|P|R|O|L|E|U|A|S|P|H|E|A|N|P|R|O|L|E|U|A|P 73  
DB 556 A|T|T|G|A|A|G|T|G|G|G|A|G|T|T|A|T|G|A|G|A|C|G|A|T|G|C|G|A|T|G|T|T|G|C|G|C|A|T|T|T|G|T|A|A 615  
QY 74 L|E|U|V|A|L|T|P|-----L|E|U|A|P|R|O|T|Y|R|T|H|L|Y|S|E|R|L|E|U|S|E|R|A|V|A|L|A|S|C|Y|S|T|P 89  
DB 616 C|G|G|C|T|G|G|G|G|A|A|G|T|T|T|A|C|A|T|G|C|G|A|A|C|A|A|T|G|T|A|T|G|A|T|A|A|G|T|T|-----T|G|T|T|G 672  
QY 90 P|H|E|Y|C|Y|M|E|R|V|A|L|E|U|A|S|P|R|O|T|Y|R|A|S|N|G|I|N|T|H|A|S|P|H|A|P|R|O|G|I|U|C|Y|S|Y|S|E|R|A|R|G 109  
DB 673 T|T|G|A|A|C|A|T|G|A|T|A|G|T|A|G|T|C|T|G|A|A|C|A|G|G|T|G|G|A|C|C|A|C|G|A|C|T|T|A|C|G|T|A|C|G|T 732  
QY 110 P|R|O|A|P|S|E|R|G|Y|L|E|U|S|E|R|A|L|E|T|H|R|G|I|U|L|E|U|A|S|P|R|O|G|I|Y|T|Y|R|I|E|T|H|R|G|I|P|R|O|L|E|U 129  
DB 733 G|C|A|G|C|A|G|G|G|C|T|T|C|G|A|T|C|A|G|T|A|T|T|T|C|A|G|C|A|G|G|G|T|A|T|----- 777  
QY 130 S|E|R|T|H|V|A|L|T|P|R|Y|G|I|U|R|P|L|E|U|Y|S|T|R|P|C|Y|V|A|I|G|U|P|H|E|G|I|Y|L|I|E|G|U|A|A|A|A| 149  
DB 778 ---T|G|A|T|T|G|A|C|A|A|A|G|T|T|T|C|A|A|A|T|C|G|G|G|A|T|A|T|G|G|C|T|A|T|G|A|A|C|C|A|T|A|A 834  
QY 150 I|L|E|V|A|L|A|V|A|P|R|O|T|Y|R|A|P|R|A|R|G|L|E|U|S|E|R|P|R|O|T|H|L|Y|S|E|R|L|E|U|G|I|U|A|P|R|E|U 169  
DB 835 A|T|G|A|C|A|G|C|G|C|G|A|T|G|A|T|T|G|A|G|C|G|T|T|G|A|T|T|A|G|A|T|T|A|G|A|A|A|C|C|G|A|T|A|G 894  
QY 170 T|Y|R|P|H|I|L|Y|L|E|U|Y|S|E|R|T|H|R|P|H|E|G|I|U|T|H|A|L|E|U|Y|S|E|R|G|Y|G|I|Y|P|R|O|S|E|R 169  
DB 895 T|A|C|T|T|A|C|G|A|A|G|C|T|A|A|A|G|A|A|C|A|A|T|C|G|A|T|T|C|A|T|A|A|T|G|G|T|G|A|A|A|A|G|T 954  
QY 190 I|L|E|V|A|P|H|E|A|H|I|S|S|E|R|E|T|G|Y|A|S|A|N|V|A|P|H|E|A|R|G|Y|R|P|H|E|U|G|I|U|R|P|L|E|U|R|G 209  
DB 955 T|G|T|T|A|T|T|G|G|A|C|A|T|T|C|A|T|G|G|G|T|T|C|A|G|A|T|T|A|T|C|T|T|T|A|C|T|T|T|A|G|A|A|T|G|G|T|C|G|A|G 1014  
QY 210 L|E|U|G|U|-----I|L|E|A|P|R|O|L|Y|S|I|S|Y|R|L|E|U|Y|S|T|R|P|L|E|U|A|S|P|G|I|N|H|I|S|I|A|L|A 227  
DB 1015 G|C|T|G|A|A|G|C|C|C|T|T|T|A|C|G|T|A|T|G|T|G|G|G|C|G|G|G|G|T|T|A|C|G|A|A|C|A|C|A|T|A|G|A|T|T|C|A 1074  
QY 228 T|Y|R|P|H|E|A|V|A|L|G|Y|A|P|R|O|L|E|U|L|E|U|G|Y|S|E|R|V|A|L|G|U|A|L|A|L|E|Y|S|E|R|T|H|L|E|U|S|E|R 247  
DB 1075 T|T|C|A|T|T|A|T|G|C|A|C|A|G|G|A|C|G|T|T|G|G|G|C|T|C|A|A|G|G|A|C|A|G|T|T|C|A|G|C|T|T|A|T|T|A|G|T 1134  
QY 248 G|Y|-----V|A|L|T|H|R|P|H|E|G|I|U|P|R|O|V|A|L|S|E|R 256  
DB 1135 G|G|T|G|A|A|T|G|A|A|G|A|T|C|A|T|T|C|A|A|T|T|A|A|T|A|C|G|T|T|A|C|C|A|T|G|T|A|T|G|T|T|G|----- 1185  
QY 257 G|U|G|Y|T|H|A|L|A|R|G|L|E|U|S|E|R|A|S|E|R|P|H|E|A|L|S|E|R|L|E|U|----- 271  
DB 1186 -----G|A|A|A|A|G|T|C|T|T|C|A|G|A|A|T|T|G|A|G|A|G|T|A|A|A|A|T|G 1221  
QY 272 -----T|P|-----L|E|U|M|E|R|P|R|O|P|H|E|R|Y|S|A|N|C|Y|S|Y|G|Y|A|S|P 283  
DB 1222 T|T|A|C|A|A|C|G|T|G|G|G|G|T|A|T|A|C|A|T|C|A|A|T|C|T|A|C|A|-----A|A|G|G|G|A|A 1286  
QY 284 A|S|N|T|H|S|E|R|T|P|R|H|I|S|P|H|E|R|G|Y|A|L|A|L|Y|S|A|P|R|Y|S|A|R|G|V|A|L|Y|R|H|I|S 303  
DB 1267 G|A|G|G|T|C|A|T|T|T|G|-----G|G|G|A|T|A|T|A|A|G|A|C|A|T|T|C|A|G|A|G|A|T|G|C|A|T|T|G|A|T 1314

```
QY 304 CysaapgluGlutylGlnSerlySerGlyTyrProThraenlleaenlle 323
   |||
   |||
   |||
Db 1315 AACACACACGACACATAC-----GGCAATTTTCATTCGATT 1350
QY 324 GlnleProserThr-----SerValThrGluThraValaenMet 338
   |||
   |||
   |||
Db 1351 GAAAGGAATACGAGGAGCTTTCACAAATAATTGACATAGAAAGCCCATTAACATG 1410
QY 339 ThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThraAlaArgGluLeuAlaap 358
   |||
   |||
   |||
Db 1411 ACA-----TTATCATATACACCTGAATAGCTCCAAAGA 1443
QY 359 GlyThrLeuPheLeuAlaIleGluLeuPheProaapSerlyPheArgMetLeuIleGln 378
   |||
   |||
   |||
Db 1444 -----AGATCATATGACGAGTACTCGTTGCGCTATTCCAAAGATGAAGAG 1491
QY 379 LeuLysLys-----LeuTyrHis-----AapAapProValPheAapProLeuThr 393
   |||
   |||
   |||
Db 1492 TTAAGAAAAAATGAGCTTACACCAAGACACTGCTGCAATCCATGGAATACCATTTCCA 1551
QY 394 ProTyrGluArgProProIleuLysAenValPheCysIleTyrGlyAlaHisLeuLysThr 413
   |||
   |||
   |||
Db 1552 -----GAGCTCCCAACATGAAA--ATCTATTGTATATACGGGGGTGAACAACCAACT 1602
QY 414 GluValGlyTyrTyrPheAlaProserGlyLysProTyrProaapAanTrpIleIleThr 433
   |||
   |||
   |||
Db 1603 GAAAGGCATAT-----1614
QY 434 AapIleIleTyrGluThrGluGlySerLeuValSerAerSerGlyThrValValAapGly 453
   |||
   |||
   |||
Db 1615 -----GTATATTAAGAAAGATGACTCTCTGCTGCTGAATTTGACATTCACATCGAA 1668
QY 454 AenAlaGlyProIle-----ThrGlyAapGluThrValProTyrHisSerLeuSer 470
   |||
   |||
   |||
Db 1669 ACAGAGCAACCTGATTTCTTCACCGAGGGGACGGAACCGTTCCGCTCGGCGCATTCGA 1728
QY 471 TrpCysLysAanTrp-----LeuGlyProLys-----ValaenlleThr 483
   |||
   |||
   |||
Db 1729 ATGTGTACAAATGGGCGCAGGCTGCTTACCGCTACACCTGCGGAATTAACGTTACT 1788
QY 484 -----MetAlaProGlnProGlu 489
   |||
   |||
   |||
Db 1789 ATTGTGGAATGAACACCAACGCAAT 1815
RESULT 13
AAS01342
ID AAS01342 standard; DNA; 1986 BP.
XX
AC AAS01342;
XX
DT 31-MAY-2001 (first entry)
XX
DE Yeast LCAT related open reading frame 1 (LRO1) gene.
XX
KW lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
KW nutritional supplement; dairy product; food product; salad dressing;
KW yeast; Arabidopsis thaliana; LCAT related open reading frame 1; LRO1; ds.
XX
OS Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT 1..1986
FT CDS /*tag= a
/*tag= a
/*product= "LRO1"
XX
XX MO200116308-A2.
XX
XX 08-MAR-2001.
XX
XX 30-AUG-2000; 2000MO-US023863.
XX
```

```
PR 30-AUG-1999; 99US-0152493P.
XX
XX (MONS ) MONSANTO CO.
XX
XX Laesener M, Van Eenennaam A;
PI
XX
XX MPI: 2001-169010/17.
XX
XX P-PSDB; AAU00464.
XX
XX New isolated nucleic acid encoding plant lecithin:cholesterol
XX acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase
XX -like polypeptides, for modifying the sterol content and oil production
XX of plants.
XX
XX Claim 5; Page 116-117; 127pp; English.
XX
XX The present sequence encoding for yeast LCAT related open reading frame,
XX (LRO1) is described in an invention relating to several novel
XX polynucleotides encoding the plant sterol acyltransferases LCAT
XX (lecithin:cholesterol acyltransferase-like; AAS001081-AAS01104, AAS01341)
XX and ACAT (acyl CoA:cholesterol acyltransferase-like; AAS01311-AAS01319).
XX A rat ACAT-like cDNA sequence (AAS01105) is also described. The
XX polynucleotides encoding LCAT or ACAT are used to produce LCAT or ACAT
XX polypeptides. They can also be used in a recombinant construct to
XX transform a host cell (preferably of a plant) or a plant. The recombinant
XX construct is used to increase or decrease the sterol content of the host
XX cell or plant. It can be used to alter oil production of the cell or
XX plant, preferably by increasing it. The oil of the plant or the plant
XX itself is used as a food product, or as nutritional or dietary
XX CC supplements, or in pharmaceutical compositions for lowering cholesterol.
XX The oil can be used in foods e.g. margarine, butter, cooking oil, and
XX CC dressings e.g. salad dressings, mayonnaise, cheese, processed meat,
XX CC pasta, sauces, cereals, desserts, dips, chips, baked goods, pastries,
XX CC cookies, snack bars, confections, chocolates, and beverages. The
XX CC alteration in sterol content and/or composition can also provide a plant
XX CC with tolerance to stress and insect damage
XX
XX
SQ Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.7e-19 Length: 1986
XX Score: 299.50 Matches: 119
XX Percent Similarity: 38.51% Conservative: 77
XX Best Local Similarity: 23.38% Miscellaneous: 188
XX Query Match: 10.10% Indels: 125
XX DB: Gaps: 24
XX
XX US-09-651-651-5 (1-546) x AAS01342 (1-1986)
QY 37 AapTyr-----SerLysLeuSerGlyIleIleIleProGlyPheAlaSerThrGln 53
   |||
   |||
   |||
Db 496 GATTATTAATATGAGGACCAACATCTCTGTGAACGTTCTCGTCAATTTTCACGGA 555
QY 54 LeuArgAlaTrpSerIleLeuAapCysAProTyrThrProLeuAapPheAapProLeuAap 73
   |||
   |||
   |||
Db 556 ATTTGAAGCTGGGAGTTATTGAGACGATGAGTCGATGTTGCGCATTTTGTGTA 615
QY 74 LeuValTrp-----LeuAapThrThrLysLeuLeuSerAlaValaenYerTrp 89
   |||
   |||
   |||
Db 616 CGGCTGTGGGAAGTTTTCATGCTGAGAACAAATGTTATGATTAAGTT--TGTGG 672
QY 90 PheLysCysMetValLeuAapProTyrAenGlnThrAapPheProGlyCysLysSerArg 109
   |||
   |||
   |||
Db 673 TTGAACATGTAATGTTAGATCTTAAACAGCTCGACCCACCAACTTTACGCTACGT 732
QY 110 ProaapSerGlyLeuSerAlaIleThrGluLeuAapProGlyTyrIleThrGlyProLeu 129
   |||
   |||
   |||
Db 733 GCAGACAGAGGCTTGAATTCAGTATTTTCATCGCAGGGAT--TTTGG 777
QY 130 SerThrValTrpLysGluTrpLeuLysTrpCysValGluPheGlyIleGluAlaAenAla 149
   |||
   |||
   |||
Db 778 --TGGAATTTGAAAGATTTTCCAAATCTGGAGATTAATTTGATGAACCAATTA 834
```

OY	150	11	levalalaval	ProTYAaRPTrAqgLeuSerProThrluVleuGIuIduAqRleu	165
		:::	:::	:::	
Db	835	ATGACGAGTGCGTGGTATGATTTGAGCGCTTCGATTTAGATTCAGAAAGCGCGATAGG			894
OY	170	TyrPhehIeVleuVleuVleuThrPheGIuThrlaIleuVleuVleuGlglyProSer			189
Db	895	TACTTATCGAAGCGTAAGAAAGAAATGMACTGTTTCATCATTAAGTAGTGGAAGAGTT			954
OY	190	11	levalPhehIaHissSerMeGIyAsnAsnValPhehIaTgTYrPheLeuGIuTrpLeuArg		209
		:::	:::	:::	
Db	955	TGTTTAAATGGACATTTCTATGGTTCACAGTATATCTTTTACTTTATGAATGGGTGGAG			1014
OY	210	LeuGIu-----11	lealProVleuHieTYrLeuVleuTYrTrpLeuAaRgIleHieAla		227
Db	1015	GCTGAAGGCCCTTTACGGTAATGATGATGGTGGTGGCTGGTTTAAAGAAACATGATTC			1074
OY	228	TyrPhehIaValGIyAlaProLeuLeuGIySerValGIuAlaIleuVleuSerThrLeuSer			247
		:::	:::	:::	
Db	1075	TTCTATTATGACGACGAGGACGCTTGGCGCTCCAAAGGAGATTCCAGCTTAATTAGT			1134
OY	248	GIy-----	ValThrPheGIyLeuProValSer		256
Db	1135	GCTGAATGAAGATACCATTCATTAATACGTTAGGCATGTATGGTTTG-----			1185
OY	257	GIuGIyThrlaIaTgLeuLeuSerAsnSerPheAlaSerSerLeu-----			271
Db	1186	-----	GAAGAATTCTTCCAAATTCAGAGAGTAAAAATG		1221
OY	272	-----Trp-----	LeuMetProPheSerIlyAsnCylyGIyAsp		283
Db	1222	TTACAAACGTGGGGTGATACCATCAATGCTATCCA-----			1266
OY	284	AsnThrSerTrpThrlaHissPheSerGIyAlaIaIaIaIyAsnRlyAaRgValTYrHis			303
Db	1267	GAGGCTCATTTG-----	GGGAGATATGAAGTATCTTCAGAGAGATGCATTGAAT		1314
OY	304	CysAaRgIuGIuGIyTYrGIuInsertyTYrSerGIyTrpProThrlaAsnIleIaAsnIle			323
Db	1315	AACAACACTGCACACTAC-----	GGCAATTTTCATTCGATTT		1350
OY	324	GIuIleProSerThr-----	SerValThrGIuThrlaIleuValAsnMet		338
Db	1351	GAAGGAAATACGACGATGCTTTCAACAAAATTGTAACATGAAGAACGCCATTAACTG			1410
OY	339	ThrSerMetGIyCyGIyLeuProThrlaLeuLeuSerPheThrlaIaArgGIuIleuAlaAsp			358
Db	1411	ACA-----	TTATCGATATACCTGAAATGCTCAAAAGA		1443
OY	359	GIyThrlaPhePheIyAlaIleGIuAaRPyIaAaRProAsnSerIlyAaRMeIeunIaGIu			378
Db	1444	-----	AGAGTACATGAGAGATACGTTCCGGCTATTCCAAAATGAAGAAAGAG		1491
OY	379	LeuIyIyIyS-----	LeuTYrHis-----	AspAaRProValPheAsnProLeuThr	393
Db	1492	TTAAGAAAATAAGACTACACACACAAAGCACTGGTCGAATCCAAATGAAGTACCACTTCCA			1551
OY	394	ProTrpArgIaAaRProProIleIyAsnAsnAlaPheCyIleIeTYrGIyAlaHissIeIyThr			413
Db	1552	-----	GAGGCTCCCAACATGA-----	ATCATTTGATATATCGGGGTGAACCCAACT	1602
OY	414	GIuValaGIyTYrTYrPheAlaAaRAsnSerGIyAsnProTYrProAaRAsnTrpIleIeIeThr			433
Db	1603	GAAGGCGCATTT-----			1614
OY	434	AspIleIeIeTYrGIuThrGIuGIySerIleuValSerAaRgSerGIyThValValaAaRGIy			453
Db	1615	-----	GTAATTAAGAGAAAGAGATGACTCTCTGCTCGAATTTGACATGACATACGAA		1668
OY	454	AsnIaIaGIyProIle-----	ThrGIyAaRGIuThrValProTYrHisSerIeIeSer		470
Db	1669	AGCAAGCAACCTGTATTTCTTCCACGAGGGGAGCGAACCCTTCCGCTGTGGCGCATTTCA			1722
OY	471	TrpCyIyAsnTrp-----	LeuGIyProIyS-----	ValAsnIleThr	483

Db 1729 ATGTGTCAAAATGGCCAGGCTTCACCGTACAACCTGCGGAAATTACGTTACT 1788  
Oy 484 -----MetaIaProGInProGlu 489  
Db 1789 ATTGTGAAATGAAACACCAAGCCAGAT 1815

RESULT 14  
ID ADF47816  
ADP47816 standard; DNA; 1986 BP.  
ADP47816;  
26-FEB-2004 (first entry)  
Yeast acyltransferase nucleotide sequence SEQ ID NO:1.

XX acyltransferase; enzyme; membrane-spanning region;  
XX active membrane independent acyltransferase; fatty acid ester;  
XX fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;  
XX phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
XX acylated amino acid; structured lipid; fat-soluble molecule; yeast; gene;  
XX ds.  
XX Saccharomyces cerevisiae.  
XX WO2003100044-A1.  
XX 04-DEC-2003.  
XX 28-MAY-2003; 2003WO-SE000870.  
XX 29-MAY-2002; 2002SE-00001581.  
XX 29-MAY-2002; 2002US-0383889P.  
XX 20-JAN-2003; 2003SE-00000142.  
XX (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.  
XX Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;  
XX WPI, 2004-053266/05.  
XX P-PSDB; ADP47818.

XX New nucleotide sequence encoding an improved acyltransferase polypeptide,  
XX useful for producing structured lipids or fat-soluble molecules, in  
XX removing undesirable fat or in modifying lipids in animal or plant raw  
XX material.  
XX  
XX Claim 5; SEQ ID NO 1; 91pp; English.

XX The present invention describes a nucleotide sequence (I) derived from a  
XX nucleotide sequence encoding an acyltransferase polypeptide comprising at  
XX least one membrane-spanning region. (I) encodes an improved active  
XX membrane independent acyltransferase polypeptide in which at least one  
XX amino acid residue of the membrane-spanning region has been deleted  
XX and/or substituted as compared to the original acyltransferase  
XX polypeptide, where the encoded active membrane independent  
XX acyltransferase polypeptide can produce fatty acid esters and/or fatty  
XX acid thioesters such as triacylglycerols, diacylglycerols,  
XX monoacylglycerols, phospholipids, glycolipids, lysolipids, waxesters,  
XX acylated carbohydrates and acylated amino acids. Also described: (1) a  
XX nucleotide sequence molecule (II) comprising at least one promoter region  
XX which functions in a host, where the promoter region is operably linked  
XX to at least one (I), which is operably linked to at least one non-  
XX translated region which functions in a host; (2) a vector comprising (II)  
XX; (3) a host cell comprising (II) or the vector; (4) producing an  
XX improved active membrane independent acyltransferase polypeptide,  
XX comprising providing the above host cell and a growth medium preparing a  
XX host cell culture, culturing the host cell culture and harvesting the  
XX host cell culture and recovering the improved active membrane independent  
XX acyltransferase polypeptide; (5) a polypeptide obtained by the method in  
XX (4), where the polypeptide at least is an improved active membrane  
XX independent acyltransferase polypeptide; (6) an oligonucleotide

specifically hybridizing to (1) under stringent conditions; and (7) a kit comprising the above polypeptide and a stabilizer. The nucleic acid molecule (1) and the polypeptide are useful in producing structured lipids or fat-soluble molecules, in removing undesirable fat or in modifying lipids present in animal and plant raw material. The present CC sequence is used in the exemplification of the present invention.

XX Sequence 1986 BP, 615 A, 361 C, 478 G, 532 T, 0 U, 0 Other;

## Alignment Scores:

Pred. No.:	1.7e-19	Length:	1986
Score:	299.50	Matches:	119
Percent Similarity:	38.51%	Conservative:	77
Best Local Similarity:	23.38%	Mismatches:	188
Query Match:	10.10%	Indels:	125
DB:	12	Gaps:	24

US-09-651-651-5 (1-546) X ADF47816 (1-1986)

```

QY 37 AepTYR-----SerlyleuSerGlyllelleleProGlyPheAlaSerThrGln 53
   |||||
DB 496 GATTATATATATGAGGCCAATCCTGTGTAATGGTTCTGTGTCTATTTCTACGGGA 555
   |||||
QY 54 LeuATGAlaTrpSerlleleuApyrProTyTrpProleuApyrPheAanProleuApy 73
   |||||
DB 556 ATGAAAGCTGGGAGACTATTGGAGACGATGAGTGGATAGTTCTGGCATTTTGGTAA 615
   |||||
QY 74 LeuValTrp-----LeuApyrThrlyleuSerAlaValAanCyTrp 89
   |||||
DB 616 CGGCTGTGGGAAAGTTTATACATGCTGAGAACAAATGATTATGATTAAGT---TGTGG 672
   |||||
QY 90 PheIysCysMetValleuApyrProTyTrpAangInThApyrIleProGluCyIysSerArg 109
   |||||
DB 673 TTGAAACATGATATGTTAATCTGTAACAGCTGAGCCACCGAACTTTACGTTACGT 732
   |||||
QY 110 ProApySerGlyleuSerAlaIleThrGluLeuApyrProGlyTyTrpIleThrGlyProleu 129
   |||||
DB 733 GCAGCAGAGGCTTCAATCACTGATTTTCATTCATGCGCAGGAT----- 777
   |||||
QY 130 SerThrValTrpIysGluTrpIleuLySTrPysValGluPheGlylleGluAlaAanAla 149
   |||||
DB 778 ---TGATTTGGAAACAAAGTTTCCAAAATCTGGAGTAATTGGCATGTAACCCCAATAA 834
   |||||
QY 150 lleValAlaValProTyTrpApyrTrpApyrSerProThrlyleuGluGluArgApyleu 169
   |||||
DB 835 ATGACGAGTGTGCGTATGATTTGGAGGCTTGCATATTAGATTAAGTAAGAACCCGATAG 894
   |||||
QY 170 TyrPheHleuIleuIysleuThrPheGluThrAlaLeuIysleuArgGlyIysProSer 189
   |||||
DB 895 TACTTACGAAAGCTAAAGCAAAATCGAACTGTTTCATCAATTTAGTGTGTAAGAAAGTT 954
   |||||
QY 190 lleValPheAlaHleuSerMetGlyAanAanValPheArgTyTrpPheleuGluTrpLeuArg 209
   |||||
DB 955 TGTATTATGACATCTTATGGGTTCTCAGATTACTTTTATCTTATGAATGGGTGAG 1014
   |||||
QY 210 LeuGlu-----lleAlaProIysHleuTyTrpIleuLyTrpLeuApyrIleHleuAla 227
   |||||
DB 1015 GCTGAAGGCCCTCTTACGTAATGAGTGTGCTGCTGGCTTAACGAAACATAGATTCA 1074
   |||||
QY 228 TyrPheAlaValGlyAlaProleuLeuGlySerValGluAlaIleIysSerThrLeuSer 247
   |||||
DB 1075 TTCATTAAATGACGACGAGGAGCTTGTGGGCTCCMAAGCAATTCACAGCTCAATTAAGT 1134
   |||||
QY 248 Gly-----ValThrPheGlyleuProValSer 256
   |||||
DB 1135 GGTGAATGAAAGATACCATTTCAATTAATACGTTAGCCATGTATGTTG----- 1185
   |||||
QY 257 GluGlyThrAlaArgleuIysleuSerAanSerPheAlaSerSerleu----- 271
   |||||
DB 1186 -----GAAAGCTTCTTCCAGAAATTAAGAGAGTAAGAAATG 1221
   |||||
QY 272 -----Trp-----LeuMetProPheSerIysAanCyIysGlyAanP 283
   |||||

```

```

DB 1222 TTACAAAGCTGGGGTGTATCATCATCACTGCTACAA-----AAGGAGAA 1266
QY 284 AanthSerTrpPheHleuPheSerGlyGlyAlaIleuIysApyrIysValIysHleu 303
   |||||
DB 1267 GAGGTCAATTTGG-----GGGATATGAAAGTATCATCTTCAAGAGATGATTTGAT 1314
QY 304 CysApyrGluGluGluTyTrpGlnSerIysTyTrpSerGlyTrpProThrAanIleIleAanIle 323
   |||||
DB 1315 AACAACTGACACATAC-----GGCAATTCATTCATTCATTT 1350
QY 324 GluIleProSerThr-----SerValThrGluThrAlaLeuValAanMet 338
   |||||
DB 1351 GAAAGCAATGACGAGCATGCTTTCAACAAATTTGACAAATGAAGAACGCCATTAACATG 1410
QY 339 ThrSerMetGlyCysGlyleuProThrIleuIysSerPheThrAlaArgGluLeuApyr 358
   |||||
DB 1411 ACA-----TTATGATATATCACTGAAATGGCTCCAAAGA 1443
QY 359 GlyThrleuPheIysAlaIleGluApyrTrpApyrProApySerIysArgMetleuHleuGln 378
   |||||
DB 1444 -----AGAGTACATGAGCATGATCTCTGCGCTATTCCAAAGATGAAGAG 1491
QY 379 LeuIysIys-----LeuTyTrpHleu-----ApyrProValPheAanProleuThr 393
   |||||
DB 1492 TTAAAGAAAATGAGCTACACCAACAGCATGTCGAATCCAAATGAAGTACCACTTCA 1551
QY 394 ProTrpGluApyrProIleIysAanValPheCysIleTyTrpGlyAlaHleuIysleuThr 413
   |||||
DB 1552 -----GAAGCTCCCATGATAA---ATCATGTATATATCGGGGTGAACAAACCAACT 1602
QY 414 GluValGlyTyTrpPheAlaProSerGlyIysPyrProApyrAanTrpIleIleThr 433
   |||||
DB 1603 GAAAGGCATAT----- 1614
QY 434 ApyrIleIleTyGluThrGluIysleuValSerIysArgSerGlyThrValAanGly 453
   |||||
DB 1615 -----GTATATTAAGAAAGAGATGATCTCTGCTGGAATTTGACATGACATACGAA 1668
QY 454 AanaIaGlyProIle-----ThrGlyApyrGluTrpValProTyTrpHleuSer 470
   |||||
DB 1669 AGCAAGCAACCTGATTTCTTCAACGAGGAGAGGAGAACCGTTCCGCTGGCGCATTTCA 1728
QY 471 TrpCysIysAanTrp-----LeuGlyProIys-----ValAanIleThr 483
   |||||
DB 1729 ATGTGTCAAAATGAGGCCAGGCTGCTTACCGTAAACCTGCGCGGAATTAAGCTTACT 1788
QY 484 -----MetAlaProGlnProGlu 489
   |||||
DB 1789 ATTTGAAATGAAACACGACGACGAT 1815

```

RESULT 15  
ADT87255 standard; DNA, 1986 BP.  
ID ADT87255 standard; DNA, 1986 BP.  
AC ADT87255;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Yeast Stress-related protein gene YNR008W.  
XX  
DE Yeast; Stress-related protein; SRP; ds; gene; environmental stress;  
KW abiotic stress; drought; heat; cold; salt.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN WO2004092398-A2.  
PD 28-OCT-2004.  
XX  
PF 15-APR-2004; 2004WO-US011888.  
XX  
PR 15-APR-2003; 2003EP-00008080.  
PR 02-MAY-2003; 2003EP-00039728.

PR 01-AUG-2003; 2003EP-00016672.  
 PR 30-SEP-2003; 2003EP-00022225.  
 XX (BADI ) BASF PLANT SCI GMBH.  
 XX Puzio P, Chardonnens A, Shirley A, Wang X, Sarria-Millan R;  
 PI Mckerlele B, Chen R;  
 XX WPI; 2004-766883/75.  
 DR P-PSDB; ADT87256.  
 PT New isolated nucleic acid molecule comprises a sequence encoding Stress-  
 PT Related Protein (SRP), useful for producing transformed plants with  
 PT altered metabolic activity resulting in increased tolerance or resistance  
 PT to environmental stress.

PS Claim 15; SEQ ID NO 233; 911pp; English.

XX The invention relates an isolated nucleic acid molecule comprises a  
 CC nucleic acid molecule encoding a Stress-Related Protein (SRP), from Yeast  
 CC or E. coli and their homologues from Rice, Soybean and Rape. Also  
 CC included are a transformed plant cell with altered metabolic activity  
 CC compared to a corresponding non-transformed wild type plant cell (where  
 CC the metabolic activity is altered by transformation with a SRP coding  
 CC nucleic acid and results in increased tolerance and/or resistance to an  
 CC environmental stress as compared to a corresponding non-transformed wild  
 CC type plant cell), a transgenic plant generated from the plant cell above  
 CC (and which is a monocot or dicot plant, or a gymnosperm plant), a seed  
 CC produced by a transgenic plant above (where the seed is genetically  
 CC homozygous for a transgene conferring altered metabolic activity  
 CC resulting in an increased tolerance to environmental stress as compared  
 CC to a corresponding non-transformed wild type plant), a nucleic acid  
 CC construct which confers the expression of the nucleic acid molecule above  
 CC (comprising one or more regulatory elements, where expression of the SRP  
 CC coding nucleic acid in a host cell results in altered metabolic activity  
 CC resulting in increased tolerance to environmental stress as compared to a  
 CC corresponding non-transformed wild type host cell), a vector comprising  
 CC the nucleic acid molecule above or the nucleic acid construct, a host  
 CC cell which has been transformed stably or transiently with the vector (or  
 CC the nucleic acid molecules above, or the nucleic acid construct), an  
 CC isolated Stress Related Protein (SRP) selected from the amino acid  
 CC sequences fully given in the specification and/or its homologues, a  
 CC method of producing a transgenic plant with altered metabolic activity  
 CC compared to a corresponding non-transformed wild type plant cell,  
 CC modifying stress tolerance of a plant, detecting environmental stress in  
 CC plant cells or plants, screening plant cells or plants for increased  
 CC tolerance and/or resistance to environmental stress, breeding plant cells  
 CC or plants towards increased tolerance and/or resistance to environmental  
 CC stress, increasing tolerance of a plant to at least one abiotic stress,  
 CC a plant transformed with the nucleic acids above and a seed of the plant.  
 CC The altered metabolic activity and/or a SRP encoding nucleic acids or its  
 CC homologues are useful as markers for selection of plants or plant cells  
 CC with increased tolerance to environmental stress, or for detection of  
 CC stress in plants or plant cells. The nucleic acids are useful for  
 CC producing transformed plants with altered metabolic activity resulting in  
 CC increased tolerance and/or resistance to an environmental stress  
 CC (drought, heat, cold and salt) as compared to a corresponding non-  
 CC transformed wild-type plant cell. The present sequence is a yeast gene  
 CC encoding an SRP of the invention.

XX Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 1,7e-19 Length: 1986  
 Score: 299.50 Matches: 119  
 Percent Similarity: 38.51% Conservative: 77  
 Best Local Similarity: 23.38% Mismatches: 188  
 Query Match: 10.10% Indels: 125  
 DB: 13 Gaps: 24

US-09-651-651-5 (1-546) x ADT87255 (1-1986)

Qy 37 AspTyr-----SerLysLeuSerGlyIleIleIleProGlyPheAlaSerThrGln 53

DB 496 GATTATATATATGAGCCAAACATCTCTGTTGAATGGTTCCTGGGTCAATTTCTACGGGA 555  
 Qy 54 LeuArgAlaTrpSerIleLeuAspCysProTyrThrProLeuAspPheAsnProLeuAsp 73  
 DB 556 ATTTGAAGCTGGGGAGTTATTTGAGACGATGAGTGAGTTCGCGCATTTTCGTA 615  
 Qy 74 LeuValTrp-----LeuAspThrThrLysLeuLeuSerAlaValAsnCysTrp 89  
 DB 616 CGGCTGTGGGAAGTTTTCATGCTGAGAACAAATGGTATGATGATTAAGTTT 672  
 Qy 90 PheLysCysMetValLeuAspProTyrAsnGlnThrAspHisAspGluCysLysSerArg 109  
 DB 673 TTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732  
 Qy 110 ProAspSerGlyLeuSerAlaIleThrGluLeuAspProGlyTyrIleThrGlyProLeu 129  
 DB 733 GCAGCAGAGGGCTTCGATCACTGATTTTCATGCGAGGGTAT----- 777  
 Qy 130 SerThrValTrpLysGluTrpLeuLysTrpCysValGluPheGlyIleGluAlaAsnAla 149  
 DB 778 ---TGATTGGAAACAAAGTTTCCAAATCTGGAGTAATGGCTATGATGATGATGAT 834  
 Qy 150 IleValAlaValProTyrAspTrpArgLysSerProThrLysLeuGluGluAspLeu 169  
 DB 835 ATGACGAGTGTGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894  
 Qy 170 TyrPheHisLysLeuLysLeuThrPheGluThrAlaLeuLysArgLysGlyProSer 189  
 DB 895 TACTTACGAGCTTAAGCAAAATCGAATCTTTCATCAATTTGAGTGAAGAAAGTT 954  
 Qy 190 IleValPheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeuGluTrpLeuAsp 209  
 DB 955 TGTTAATGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1014  
 Qy 210 LeuGlu-----IleAlaProLysHisTyrLysLysTrpLeuAspGlnHisIleAla 227  
 DB 1015 GTTGAAGGCCCTTTTACCGTAAATGGTGTGCTGGCTGGGTAAACAAACATAGATTCA 1074  
 Qy 228 TyrPheAlaValGluAlaProLeuLeuGlySerValGluAlaIleLysSerThrLeuSer 247  
 DB 1075 TTGATTATATGACACAGGAGGACGCTTGGGGGCTCCAAAGCAGATTCAGCTTAATTAGT 1134  
 Qy 248 Gly-----ValThrPheGlyLeuProValSer 256  
 DB 1135 GGTGAATGAAAGATACCATTCATTAAATACCTTACCCATGATGATGATGATGATGAT 1185  
 Qy 257 GluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeu----- 271  
 DB 1186 -----GAAAGCTTCTTCTCAAGATTTGAGAGAGTAAAAATG 1221  
 Qy 272 -----Trp-----LeuMetProPheSerLysAsnCysLysGlyAsp 283  
 DB 1222 TTACAAACGTTGGGTGATATACATCAATCTACCA-----AAGGGGAA 1266  
 Qy 284 AsnThrSerTrpThrHisPheSerGlyAlaAlaLysLysAspLysValTyrHis 303  
 DB 1267 GAGGTCAATTTGG-----GGGATATGATGATCATCTTCAGAGATGATGATGAT 1314  
 Qy 304 CysAspGluGluGluLysTrpGlnSerLysTyrSerGlyTrpProThrAsnIleLeuAsnIle 323  
 DB 1315 AACCAACATGACACATAC-----GGCAATTTCAATTCATTCATT 1350  
 Qy 324 GluIleProSerThr-----SerValThrGluThrAlaLeuValAsnMet 338  
 DB 1351 GAAAGAAATACGAGCGATGCTTTCACAAATAATTTGACATGAAAGCGCATTAACATG 1410  
 Qy 339 ThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAsp 358  
 DB 1411 ACA-----TTATGATATACACCTGATGATGATGATGATGATGATGATGATGAT 1443  
 Qy 359 GLyThrLeuPheLysAlaIleGluAspTyrAspProAspSerLysArgMetLeuHisGln 378

```
Db      1444 -----AGAGTACATGAGCATCTCGCTATTCCAGAAATGAGAAAG 1491
Qy      379 LeuLysLys-----LeuTyrHis-----AspAspProValPheAsnProLeuThr 393
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1492 TTAAGAAAAAATGAGCTACACCAAGCACTGATCGAATCAATGAGAACTACCACTTCCA 1551
Qy      394 ProTPGluArgProProIleLysAsnValPheCysIleTyrGlyAlaHisLeuLysThr 413
          1552 -----GAAGCTCCCAATGAAA--ATCTATGTATATACGGGGTGAACAACCACT 1602
Qy      414 GluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAsnTrpIleIleThr 433
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1603 GAAAGGGCATAT----- 1614
Qy      434 AspIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyThrValValAspGly 453
          ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1615 -----GTATATAGAGAGAGATGACTCTCTCTGATGAATTGACCATCGACTACGAA 1668
Qy      454 AsnAlaGlyProIle-----ThrGlyAspGluThrValProTyrHisSerLeuSer 470
          ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1669 AGCAAGCAACTGTATTTCTCTACCCAGAGGGGAGAGGAACTTCCTCGTGGCGCATTTCA 1728
Qy      471 TrpCysLysAsnTrp-----LeuGlyProLys---ValAsnIleThr 483
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1729 ATGTGTCACAAAATGGGCCAGGGTGTCTCACCGTACACCTGCGCGAATTAACGTTACT 1788
Qy      484 -----MetAlaProGlnProGlu 489
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1789 ATTGTGAAATGAAACCAAGCCAGAT 1815
```

Search completed: November 22, 2005, 11:10:37  
Job time : 679 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 22, 2005, 10:51:45 ; Search time 208 Seconds  
(without alignment)

4666.099 Million cell updates/sec

Title: US-09-651-651-5

Perfect score: 2965  
Sequence: 1 MGANSKSVTASFTVIAVFLL.....DSESIPTKRTAVWELDKSGY 546

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h  
-O/cgcn2\_1/USPTO.spool/US09651651/runat.18112005.101831.20277/app.query.fasta\_1.711  
-DB=Issued\_Patents\_NA -QFMT=fasta -SUFFIX=rm -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=biosum62 -TRANS=human0.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09651651 @CGCN 1.1 143 @runat.18112005.101831.20277 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEJOLERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DEBEXT=7

Database :

Issued Patents NA: \*  
1: /cgcn2\_6/ptodata/1/ina/1 COMB.seq: \*  
2: /cgcn2\_6/ptodata/1/ina/5 COMB.seq: \*  
3: /cgcn2\_6/ptodata/1/ina/6A COMB.seq: \*  
4: /cgcn2\_6/ptodata/1/ina/6B COMB.seq: \*  
5: /cgcn2\_6/ptodata/1/ina/H COMB.seq: \*  
6: /cgcn2\_6/ptodata/1/ina/PCUS COMB.seq: \*  
7: /cgcn2\_6/ptodata/1/ina/PP COMB.seq: \*  
8: /cgcn2\_6/ptodata/1/ina/RE COMB.seq: \*  
9: /cgcn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	10.9	2106	3	US-09-248-796A-1147
2	234	7.9	1137	3	US-09-402-532-14
3	234	7.9	1236	3	US-09-402-532-19
4	226	7.6	1354	3	US-09-949-016-5703
5	226	7.6	1744	3	US-09-054-272-33
6	226	7.6	1744	3	US-09-919-497-30
7	220.5	7.4	1137	3	US-09-402-532-12
8	220.5	7.4	1146	3	US-09-402-532-15
9	220.5	7.4	1236	3	US-09-402-532-17

10	220.5	7.4	2680	3	US-09-999-833A-156	Sequence 156, App
11	220.5	7.4	2680	3	US-10-020-445A-156	Sequence 156, App
12	210	7.1	2687	3	US-09-489-847-57	Sequence 57, App1
13	209	7.0	1233	3	US-09-402-532-13	Sequence 13, App1
14	209	7.0	1242	3	US-09-402-532-16	Sequence 16, App1
15	209	7.0	1332	3	US-09-402-532-18	Sequence 18, App1
16	147	5.0	1332	3	US-09-857-612A-13	Sequence 13, App1
17	144	4.9	1500	3	US-09-857-612A-13	Sequence 9, App1
18	128	4.3	1117	3	US-09-857-612A-5	Sequence 5, App1
19	120	4.0	8230	3	US-09-949-016-15445	Sequence 15445, A
20	117	3.9	1440	3	US-09-857-612A-7	Sequence 7, App1
21	109	3.7	6696	3	US-10-272-459-29	Sequence 29, App1
22	108	3.6	1956	3	US-10-272-459-39	Sequence 39, App1
23	108	3.6	6757	3	US-10-272-459-30	Sequence 30, App1
24	108	3.6	7478	3	US-10-104-966-15	Sequence 15, App1
25	108	3.6	7478	3	US-09-929-955-15	Sequence 15, App1
26	107	3.6	10718	3	US-08-325-426B-1	Sequence 1, App1
27	106.5	3.6	2512	3	US-09-702-705-1802	Sequence 1802, Ap
28	106.5	3.6	2512	3	US-09-736-457-1802	Sequence 1802, Ap
29	106.5	3.6	2512	3	US-09-671-325-1802	Sequence 1802, Ap
30	106.5	3.6	2512	3	US-10-017-754-1802	Sequence 5811, Ap
31	104.5	3.5	3673	3	US-09-949-016-5811	Sequence 4772, Ap
32	104	3.5	5120	3	US-09-949-016-4772	Sequence 3, App1
33	104	3.5	12381	3	US-09-953-096-3	Sequence 1, App1
34	104	3.5	16557	3	US-09-953-096-1	Sequence 8210, Ap
35	103	3.5	2240	3	US-09-902-540-8210	Sequence 847, App
36	103	3.5	8060	3	US-09-902-540-847	Sequence 18, App1
37	102.5	3.5	1442	2	US-08-468-763-18	Sequence 6417, Ap
38	102.5	3.5	1442	2	US-08-393-996A-18	Sequence 2002, Ap
39	102.5	3.5	2856	3	US-09-248-796A-6417	Sequence 3, App1
40	101.5	3.4	3579	3	US-09-949-016-2002	Sequence 3, App1
41	101	3.4	7488	3	US-08-475-886-3	Sequence 3, App1
42	101	3.4	7488	3	US-09-653-499-3	Sequence 3, App1
43	101	3.4	7488	3	US-10-135-988-3	Sequence 1, App1
44	101	3.4	7493	3	US-08-475-886-1	Sequence 1, App1
45	101	3.4	7493	3	US-08-397-232-1	Sequence 1, App1

#### ALIGNMENTS

RESULT 1  
US-09-248-796A-1147  
; Sequence 1147, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248, 796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074, 725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096, 409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 1147  
; LENGTH: 2106  
; TYPE: DNA  
; ORGANISM: Candida albicans  
; NAME/KEY: unsure  
; LOCATION: (341)  
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown.  
US-09-248-796A-1147  
Alignment Scores:  
Pred. No.: 7.91e-28  
Score: 324.00  
Percent Similarity: 38.29%  
Best Local Similarity: 24.00%  
Query Match: 10.93%  
DB: 3  
Length: 2106  
Matches: 126  
Conservative: 75  
Mismatch: 214  
Indels: 110  
Gaps: 21

```

US-09-651-651-5 (1-546) x US-09-248-796A-1147 (1-2106)
QY      28 VALGIUASPGLUTHRGLUPHEHISGLYASPTYSERLYLEUSERGILYILEILEIPEPRO 47
      596 CTTGAAATTCGAAAAATTTTCAACAGCTGACCAATATGTA-----GTATNGCCCT 636
QY      48 GILYHEALASERTHGLINLEUARGALATPSERIILEU-----ASRCYSPPTYR 64
      637 GGGGTGATTTCATCGGTTAGAAATCTTGGGGACCCACCACTACGCGGGATTGTCATCA 696
QY      65 THRPROLEU-----ASPPHEANPROLEUAPHEUVALTRPLEU 77
      697 ATTGCGCATTTCCGAAAAAGATTATGCGGATCATTTTATATGTTAAAGACTATGCTTTTA 756
QY      78 ASPRTHRLHYLEULEUSERALAVALEANCYSITRPHELYCYMERVALLEUAPRO 97
      757 GATTAATACT-----TGTGGCTTAGACATATCATGTTAGATACC 795
QY      98 TYRANGINTHRASPHISPROGLUCYLYSERARGPROASPSERGLYLEUSERALALE 117
      796 ACTACTGATTAATGATCCTCTAATATTAAGTTAGACAGCTCAAGGTTGAAGCCGCT 855
QY      118 THRGLULEUASPARGIYTYRILETHRGILYPROLEUSERTHRVALTRPLYSGLUTRPLEU 137
      856 GATTTTTCATGCGGGGTAAT-----TGGAATTGGAAATAAATTTTA 897
QY      138 LYSTRCYVALGLUPHEGLYILEGLUALAANALALEVALALAVAPROTYRABPTP 157
      898 CAAACTTAGCTGTGATGTTATGGAACCCCAATTAATGATTATAGTCCGTATGACTCG 957
QY      158 ARGLEUSERPROTHRYLYLEUGLUARGAPLEUTRPHENILYLYSEULEUYSLEUTHR 177
      958 ACATTAACCTTAATATGATTAGAGAAAAGAGATGATATTTCTTAATATGAAAGCAAA 1017
QY      178 PHEGLUTHRALALEUYSLEUARGLYGLYPROSERILEVALPHEALANISERMEGLY 197
      1018 GTTGAATATGTCAAACAAATTGACAGCTAAAAATCTGATTATGCGGCATTCAAAGGCT 1077
QY      198 AANAENVALPHEARGTYRPHLEUGLUTRPLEUARGLEUGLULEALPROLYSHISTYR 217
      1078 TCACAATATTATTAATATTTTGAATG-----GTTGAAGCCAAAGAGAAATATTAAT 1131
QY      218 -----LEUYSITRLEUARGLINHISILEALTYRPHLEALVALGLYALA 233
      1132 GGTAACGAGACCTTAATGGGTGCGAAGATTAGCGTTGAACGATTGTAATATTAGTGA 1191
QY      234 PROLEULEUGLYSERVALGLUALALEYSERTHRLEUSERGLY-----248
      1192 TCTTCTTTGGTACACCAAGGCTATCCAGCATTAATTTCTGTGAATAAGTAAGTACA 1251
QY      249 -----VALTHRPHGLYLEUPRO-----VALSERGLUYTHRALA 260
      1252 GTACAAATTAATGCTTACCGGTGCTGATGATTAGACCAATTTTTCAGTAGACGTGAAGA 1311
QY      261 ARGLEULEUSERANSERPHEALASERSEULEUTRPLEUMECPROPHESERYSANCYVS 280
      1312 GTCGATATGTTGAGATCATTTGGTGTATTTGCCAGTATGATACCT-----1356
QY      281 LYSGLYASPDANHTSERTRTHRHSIPHESERGILYVALALALYSLYASPLYSARG 300
      1357 AAAAGGTGTCACAAATATTCGGGCAATTGACT-----TATGCTCTGATGATGAATA 1410
QY      301 VAL--TYRHIACYAPGLUGLUGLUTYRGLINSELYTYSERGLYTRPROTHRAEN 319
      1411 GTTTCATTTGATACCGAAAAAGAAATATTTGTAAGAAAAAGATCATTT-----1461
QY      320 ILEILEANLEULLEPROSERTHRSEVALTHRGILUTHRLALEUVALANMETTHR 339
      1462 -----GGATCATTTTATTCAAATACAAACCTGCCAATGACTCT 1497
QY      340 SERMEGLUCYSEGLYEUPEUTHRLEUPEUTHRLEUUSERPHETHRILAARGLULEUVALAERGLY 359

```

```

DB      1498 TCTAGAGAAAGTAACTATTGACCAAAAGTATTGAAGATTTATGAAAAATTTCCAGACTGG 1557
QY      360 THRLEUPHELYVALALEGLUASPTYTRAPROASPSERYVALGMELEUHIIGLILEU 379
      1558 TATTCCAAAGACGCGGAAAAATTTATTCCTTGGCGCTGCTCAGACAAAGAAAGATTG 1617
QY      380 LYSLYSEULEUTYRHIASAPRPROVALPHEANPROLEUTHRPROTRPGLUARGPRO--- 398
      1618 GAAAGAAATATCATGATCAA-----CGAATAGCTTAATCCATTG 1659
QY      399 -----PROILEYSANVALPHECYSLIETRYGLYALAHISLEU 411
      1660 GAGCGAGATTACCAATATGCTCAAGTTTGAAAGTTTATGTTTATGCGTTGTAAC 1719
QY      412 LYSTHRGLUVALGLUTYRTRPHLEALAPROSERGLYSPROYRPROASANTRPLE 431
      1720 CCTACTGAAGAGGCTTAATAATACATGCTCTGTATTAAGTCAACAAACCTTGATTATGTA 1779
QY      432 ILETHRAPILIEILETYRGLUTHRGILUYSERLEUVALISERARGSERGLYTHRVALVAL 451
      1780 ATT-----GATGCTGATTTCTCCTGATGCTGTGTTCTT 1812
QY      452 ASPGLYANALAGLYPROLIETHRGLYAPGLUTHRVALPROTYRHSERLEUSERTRP 471
      1813 GGT-----GATGGGATGGACAGTTTCATTGTTAAACCATACATG 1854
QY      472 CYSLYASANTRP--LEUGLYPROLYS-----VALASNIETHR 483
      1855 TGTCTAGATGACCAAAAGGTGATTAATCAAGATACAACTCTGTAATGTCAAATGTCACC 1914
QY      484 METALAPRO---GLNPROGLINHISAPGLYSEAPRVAL-----495
      1915 ATGTGTAATTAATTAACATCAACCAACCGATGATTGATTTAAGAGTGGAGCTAAACCTGCT 1974
QY      496 ---HISVALGLILEU 499
      1975 GAACACGTTGATATT 1989
DB

```

```

RESULT 2
US-09-402-532-14
; Sequence 14, Application US/09402532
; Patent No. 6498019
; GENERAL INFORMATION:
; APPLICANT: Taniyama, Yoshio
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
; FILE REFERENCE: 2456USOP
; CURRENT APPLICATION NUMBER: US/09/402,532
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/JP98/01643
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: JP 10-010289
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: JP 9-184885
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 9-093355
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE:
; SEQ ID NO 14
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Murine
US-09-402-532-14

Alignment Scores:
Pred. No.: 1,596-17 Length: 1137
Score: 234.00 Matches: 116
Percent Similarity: 25.81% Conservative: 67
Best Local Similarity: 22.70% Mismatches: 154
Query Match: 7.89% Indels: 175
DB: 3 Gaps: 26

US-09-651-651-5 (1-546) x US-09-402-532-14 (1-1137)

```



```

Db      175 AAGTTGTAAGTACTTGTCTCAAGAGACGACGACTACTTACA----- 222
Qy      75 ValTrpLeuAerThrThrLeuLeuSerAla---ValAsnCysTrpPheLeuCysMet 93
Db      223 CTCTGGCTGATGCTGGAAGTACTGCTGCTGTTATCATGACTGCTGATGACAAATATC 282
Qy      94 ValLeuAerProTyrAlaGlnThrAspHisProGluCysLeuSerArgProAspSer--- 112
Db      283 AGGCTG---GTTTACCAACAGAACATCT-----CGGGCCACCCAGTTTCCCATGATGTG 333
Qy      113 -----GlyLeuSer-----AlaIleThrGluLeuAerProGly--- 123
Db      334 GACGTCGCTGCTCCCTGGCTTTGGGAAACATTTTCTATGGAATCTCTAGACCCCAAG 393
Qy      124 TyrIleThrGly-ProLeuSerThrValTyrPlyGluTrpLeuLeuSerProCysValGluPhe 143
Db      394 AGGAATGTGGTCTTCACTATTCTTACACTATGTG-GAGAGCCTT-----GTGGGCTG 443
Qy      143 eGly---IleGluAlaAsnAlaIleValAlaValProTyrAerTrpArgLeuSerProThr 162
Db      444 GGGCTACACACGCGGTGAAGACGTTGAGGTGCTCCATAGATGGCGGCGAGCCCA-- 501
Qy      162 rLeuLeuGluGluAerLeuTyrPheHisLeuLeuLeuThrPheGluThrAlaIle 182
Db      502 -----AATGAAACGCGCCCTACTTCTGGCCCTGCGAGATGATGAGAGATGTA 554
Qy      182 uLeuLeuAerGluGlyProSerIleValPheAlaHisSerMetGlyAsnValPheArg 202
Db      555 CAGAGTGTATGGGGGGCCCGGTGCTGCTGCGCCACACATGGGCAAGTGTACATGCT 614
Qy      202 gTyrPheLeuGluTrpLeuAerGluLeuAlaPheArgHisTyrLeuLeuTrpLeuAs 222
Db      615 CTACTTCTGCGAG-----CGGCAGCCCAAGTCTGGAAGGA 650
Qy      222 pClnHisIleHisAlaTyrPheAlaValAlaGlyAerProLeuLeuGlySerValGluAla 242
Db      651 CAATATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710
Qy      242 eLysSerThrLeuSerGlyValThrPheGlyLeuProVal---SerGluGlyThrAlaArg 261
Db      711 GCGTGCCTGCGCTCAGAGACACAAATCCCATTTCCCGCATTTGGCGCATGAAGATCCG 770
Qy      261 gLeuLeuSerAerAerPheAlaSerSerLeuTrpLeuMetProPheSerLeuAsnCysVal 281
Db      771 GGAACGACGCGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Qy      281 eGlyAerAerThrSerTrpThrHisIerPheSerGlyAlaAlaLeuLeuAerLeuArgVal 301
Db      817 -----AACCAACATTTGCTGAT-----GAAAGAGT 842
Qy      301 lTyrHisCysAerGluGluGluTyrGlnSerLeuTyrSerGlyTrpProThrAsnIle 321
Db      843 ATTT----- 846
Qy      321 eAsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsnMetThrSerMet 341
Db      847 ----GATATCAGACGCGCATGACTAGTACAG----- 873
Qy      341 tGluCysGluLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAerProTyrHis 361
Db      873 ----- 873
Qy      361 uPheLeuValAlaIleGluAerTyrAerProAerSerLeuValGMetLeuHisGln---Leu 380
Db      874 -----CTCCGGGAGCTAT-----CACCGGCTTCTTCCG 899
Qy      380 eLysLeuTyrHisAerAerProValPhe-----AsnProLeuThrProTrp 395
Db      900 GGAACATCGGCTTGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
Qy      395 pGluAerProProIleLeuAerValPheCysValTyrGluAlaHisLeuLeuThrGluVal 415

```

```

Db      960 CATAGCGCCACCGGGGTGAGCTGACTGCTTGATGAGCAGTGGTGTCCACGCCAA 1019
Qy      415 lGlyTyrTyrPheAlaProSerGlyLeuProTyrProAerAsnTrpIleIleThrAerP 435
Db      1020 CTCTTCTCTACTAC-----GAGAGCTTTCCTGAT----- 1047
Qy      435 eIleTyrGluThrGluGlySerLeuValSerArgSerGlyThrValValAerGlyAsnAl 455
Db      1048 -----CGGAGCCCAAAATCTGCTTCCGGGAT-- 1074
Qy      455 aGlyProIleThrGlyAerGluThrValProTyrHisSerLeuSerTrpCysValAerTr 475
Db      1075 -----GGTACGCGCAGCTGAACTGGAGAGCGTCTCTGACAGTCCCAAGCCTG 1121
Qy      475 pLeuGly-----ProLysValAsnIleThrMetAlaProGlnProGluHisAerP 492
Db      1122 GCAGAGCGCCCAAGACACAGATATCTTCAGAGAGTCCCGGAGACGAGAC----- 1176
Qy      492 ySerAerValHisValGluLeuAerValAerHisGluHisGlySerAerPheIleAlaAs 512
Db      1177 -----ATTGAGATGTAGCCAA 1193
Qy      512 pMetThrLysAlaProArgValLysTyrIle 522
Db      1194 TGCCACGACCTTGCTGCTTATCGAAACGTGTG 1224

```

## RESULT 4

```

US-09-949-016-3703
/ Sequence 3703, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3703
/ LENGTH: 1354
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-3703

```

```

Alignment Scores:
Pred. No.: 1,946-16 Length: 1354
Score: 226.00 Matches: 114
Percent Similarity: 35.82% Conservative: 54
Best Local Similarity: 24.31% Mismatches: 153
Query Match: 7.62% Indels: 148
DB: 3 Gaps: 25

```

```

US-09-651-651-5 (1-546) x US-09-949-016-3703 (1-1354)
Qy      44 lIleIleProGlyPheAlaSerThrGlnLeuAerGluAla----- 56
Db      159 ATCTCTGTCGCGCGCTGCTGCGGAAATGACGTAAGAACGGAAGCTGAGACAAACAGATG 218
Qy      57 -----TrpSerIleLeuAerCysAerProTyrThrProLeuAerPheAerProLeuAerLeu 74
Db      219 GTGAATGGATG-----TGTAACGCAAGACAGAGACTTC-----TTCAAC 260
Qy      75 ValTrpLeuAerThrThrLeuLeuLeuSer---AlaValAsnCysTrpPheLeuCysMet 93
Db      261 ATCTGCTGAGATGCAACATGTTCTTACCCCTTGGGATGAGACTGTG----- 308

```



```

QY 44 ILEIIELEPROGLYPHEALASERTHRG1NLEUVALA----- 56
DB 415 ATCTCTGTCGCCGCTGCGGGAATCAGTGAAGCAAGCTGGACAACAGATGTG 474
QY 57 -----TTPSERILEUABPCYAPROTYRTHRPROLEUABPHEANBPROLEUABP 74
DB 475 GTGAAGTGGATG-----TGCTACCGGCAAGACAGAGAGACTTC-----TTCACC 516
QY 75 VALITRPLEUABPTHRTHRILYBUELEUSER---ALAVALAENCYATRPHELYBCEWET 93
DB 511 ATCTGGCTGGATCTCAACATGTTCTTACCCCTTGGGTAGACTGTGG----- 564
QY 94 VALLEUABP-----PROTYRANGIN-----THRAPHISPRO 104
DB 565 ---ATCGATTAACACACAGGTTGTCTTACAACCGGAGCTGTGGCTCGTTCACAGCCCT 621
QY 105 GLUCYBLYSERARG---PROABPSERG1LYLEUSER---ALAIETHRG1LEUABP 122
DB 622 GGTGTCCAGATCCGCTGCTTGGCTTGGCAAGACCTACTCTGTGAGTACTGGACAC 681
QY 123 GLIYTRILETHRGLYPROLEUSERTHRVAITRPLYEGITRPLEUABPTRPCYVALG 142
DB 682 ACAGAGCTGGACGGTACTGGACACACTG-----GTGCAGAACCTGTCTCAAC 729
QY 143 PHEGLY---ILEGLUALABNALALEVALAVALAVALAVALAVALAVALAVALAVALA 161
DB 730 AATGGCTTACGTGGGAGGACAGACTGTGGCGCGCGCCCTATGATCTGGCTGAGGCC 789
QY 162 THRILYLEUG1UG1UDRGAAPLEUTYRPHNISLYLEUABPHEUTHRHEGUTR 181
DB 790 GGCACGACGAGAGAG-----TACTACCGCAAGCTGGCAAGGCTGGTGGAGAGAG 840
QY 182 LEULYLEUARG1LYGLYPROSERILEVALPHEALNHSERNEGLYABNANVALPHE 201
DB 841 CACGCTGCTATGGGAAGCTGTCTTCTCATTTGACACAGCTCGCTGTACTACACT 900
QY 202 ARGTYRPHLEUG1UTRPLEUABPGLU1LEALPROLYBHSITRYLEUABP 221
DB 901 CTCTATTTCTCTG-----CTGCCCAAGCCCAAGGCC-----TGGAG 936
QY 222 AERGLNHSI1LEI1VALATYRPHALVALAVALAVALAVALAVALAVALAVALAVALA 241
DB 937 GACCGCTTATGATGCTTCACTCTCTTGGGCTCCCGGGGTGGCTCCATCAAGCC 996
QY 242 ILELYSERTHRLEUSERGLYVALITHRPHEGLYLEUPROVALSERGLU1THR 261
DB 997 ATCTGCTGTGGCTCAGGTGACAACCAAGGCGATCCCATC----- 1038
QY 262 LEULEUSERABNBSERHEALASERTLEUTRPLEUMERPROPHESERLYBANCY 281
DB 1039 -----ATGTCAGCATCAAGCTG----- 1056
QY 282 GLYABRABNBSERTRTHRNPHEBSERG1LYALALALYBLYBAPRYBARGVAL 301
DB 1057 -----AAAGAGAGAGAGCCGAT 1074
QY 302 TYRNPISYABRGLUG1UG1UTRGLNBSERYLYRSERG1YTRPRTHRANLE 321
DB 1075 ACCACACACCTCCCTGATGTTCTTCCCTCTGCATGGCG---TGGCTGAGACCA 1131
QY 322 ANPILLEG1ULERPROSETHRSERVALTHRG1UTHRALALEUVALABNBSER 341
DB 1132 TTCATTTCCACACCCAGC----- 1149
QY 342 GLUCYBGLYLEUPROTHRILEUABPHEALARG1LEUVALAABRGLYTHR 361
DB 1150 -----TTCACACTACACAGCGCT-----GACTCTCAACCG 1179
QY 362 PHELYALALLEG1UDRPTYRABPPOABRSERLYBARGHELEUHSI1LEUABP 381
DB 1180 TTCTTTGACAGACTGCACTTTGAG---GAAGGCTGTGATCATGTGGCTGCAGTCA 1236

```

```

QY 382 LEUTYRNPISABRABPPOVALPHEANBPROLEUTHRPROTRGLIABXPROPRO 401
DB 1237 CTCTGGACGAGACTCCACGA-----CTGGGTG 1266
QY 402 ANVALPHECYBIETRYG1YALAHILEUABPHEUTHRGLUVALGLYTRYRPH 421
DB 1267 GAAGTATGCTGTCTTACGCGGTGGCTGCCACAGCCCGACCTATCATCTAC---GAC 1323
QY 422 SERGLYBSPROTYRPROABRABNTRPILETHRABPILIEITYGLUTHRO 441
DB 1324 CACGCTTCCCTTACACGAGCTT-----GTGGTGTGCTCTATGAG----- 1365
QY 442 SERLEUABPSEARGSERGLYTHRVALAABRGLYABNALAGLYPROILETHR 461
DB 1366 -----GATGTGAT 1374
QY 462 GLUTHRVALPROTYRNPISERLEUSERTRPCYBLYBANTRPLEUG1PROLY 481
DB 1375 GACACGGTGGGACCCGACGACCGAGCTGTGGCTGTGGCAGGGCGCGCAG----- 1428
QY 482 ILETHRMEALAPROGLINPROGLINHIS 490
DB 1429 -----CCACAGCTGTGCAC 1443

```

```

RESULT 6
US-09-919-497-30
; Sequence 30, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Multec, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-30

```

```

Alignment Scores:
Pred. No.: 2,996-16 Length: 1744
Score: 226.00 Matches: 114
Percent Similarity: 35.82% Conservative: 54
Best Local Similarity: 24.31% Mismatches: 153
Query Match: 7.62% Indels: 148
DB: 3 Gaps: 25

```

```

US-09-651-651-5 (1-546) x US-09-919-497-30 (1-1744)
QY 44 ILEIIELEPROGLYPHEALASERTHRG1NLEUVALA----- 56
DB 415 ATCTCTGTCGCCGCTGCGGGAATCAGTGAAGCAAGCTGGACAACAGATGTG 474
QY 57 -----TTPSERILEUABPCYAPROTYRTHRPROLEUABPHEANBPROLEUABP 74
DB 475 GTGAAGTGGATG-----TGCTACCGGCAAGACAGAGAGACTTC-----TTCACC 516
QY 75 VALITRPLEUABPTHRTHRILYBUELEUSER---ALAVALAENCYATRPHELYBCEWET 93
DB 511 ATCTGGCTGGATCTCAACATGTTCTTACCCCTTGGGTAGACTGTGG----- 564
QY 94 VALLEUABP-----PROTYRANGIN-----THRAPHISPRO 104
DB 565 ---ATCGATTAACACACAGGTTGTCTTACAACCGGAGCTGTGGCTCGTTCACAGCCCT 621
QY 105 GLUCYBLYSERARG---PROABPSERG1LYLEUSER---ALAIETHRG1LEUABP 122
DB 622 GGTGTCCAGATCCGCTGCTTGGCTTGGCAAGACCTACTCTGTGAGTACTGGACAC 681

```





```

Db      |||||.....|||||
465  TGGGGCCCGTGGTGTCTGTTGCCACAGTATGGCAACATGACGCTTACTTTCT 524
Qy      |||||
205  UGUUUTPLUeAUGLeUgLUlLeAlAProLyshIeUlyeUTrPLUeApgLnHlE1 225
Db      |||||
525  GCAG-----CGGACGCCGACGGCTGGAAAGCAAGATATAT 560
Qy      |||||
225  eHlSaLaTyRPhEAlAValGlyAlAProLeUeUgLySeRValGluAlIleLySeRTh 245
Db      |||||
561  CCGGGCTTCGTGTCACTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCT 620
Qy      |||||
245  rLeUSeRgLyValThRPhEgLyLeUProVal---SeRgLUgLyThAlAArgLeUSe 264
Db      |||||
621  GCGTTGAGAGACACACACCGGATCCAGTATCGGCGGCGGCGGCGGCGGAGAGAGA 680
Qy      |||||
264  rAnSeRPhEAlASeRSeRLeUTrPLUeMeRProPhSeRlySeAnCyAlSeGlyASeR 284
Db      |||||
681  GGGGTCAAGTGTCTCCACACAGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 719
Qy      |||||
284  nThSeRTrPrThRPhESeRgLyGlyAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 304
Db      |||||
720  CTACACATGATCA----- 732
Qy      |||||
304  sASeRgLUgLUgLyUTrgInSeRlySeRgLyUTrProThRAnIleIleAnIleG1 324
Db      |||||
732  ----- 732
Qy      |||||
324  uIlEProSeRThSeRValThRgLUThRAlAlAlAlAlAlAlAlAlAlAlAlAlAl 344
Db      |||||
733  ---CCTGAGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 774
Qy      |||||
344  yLeUProThRLeUSeRPhEThRAlAArgLUlLeUAlAlASeRgLyThLeUPhElyAl 364
Db      |||||
774  ----- 774
Qy      |||||
364  aIlEgLUaSeRgLyASeRProASeRSeRlySeAArgMeLUhISeLUlySeLUyRnH 384
Db      |||||
775  -CTGCGGGACTAC-----CGCAAGTCTTCCA 800
Qy      |||||
384  sASeR-----ASeRProValPhEAnProLeUTh 393
Db      |||||
801  GGACATCGGCTTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
Qy      |||||
393  rProTrPrGLuASeRProProIleLySeAnValPhESeRleTyRgLyAlAlIleUy 413
Db      |||||
855  GGAAGCCAGATGCCACTGGCGTGCAGCTGCAGCTGCCTTATGCG----- 900
Qy      |||||
413  rGluValGlyTyTrPhEAlAProSeRgLyLySeProTyRProASeRAnTrIleIle 433
Db      |||||
901  -----ACTGGCTGCCCAACACAGACTCTTC----- 927
Qy      |||||
433  rASeRleIleTyRgLUThRgLUgLySeRLeUValSeAArgSeRgLyThRValAlASe 453
Db      |||||
928  -----TACTATGAG-----AGCTTCCTGACCGGACCCATAAATCTGCTTGG 971
Qy      |||||
453  yASeNAIaGlyProIleThRgLyASeRgLUThRValProTyRnHSeRLeUSeRPro 473
Db      |||||
972  TGAC-----GGCATGTACTGTGAATTGAAGAGTGCCTGCAGTGCAC 1016
Qy      |||||
473  sAnTrPLUeUgLyProLySeValASeNleThMeAlAProGLnProGLnHlASeRgLy 493
Db      |||||
1017  GCGCTGCGAGAGCCGCGCAGAGACCAAGTGTGCG-----CAGAGTGCAGAGGAG 1070
Qy      |||||
493  rASeRValHlAValGlyLeUSeAnValASeRnHlASeRgLUhISeRAnTrIleIle 513
Db      |||||
1071  CGAG-----CACATGAGATG-----CTGGCCAAACG 1097
Qy      |||||
513  tThRyAlAProAArgValLyTyRle 522
Db      |||||
1098  CACCAACCTGGCTTATCTGAACGTGTG 1125

```

```

; Sequence 15, Application US/09402532
; Patent No. 6498019
; GENERAL INFORMATION:
; APPLICANT: Taniyama, Yoshio
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
; FILE REFERENCE: 245USOP
; CURRENT APPLICATION NUMBER: US/09/402,532
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: JP 10-010289
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: JP 9-184885
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 9-093355
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE:
; SEQ ID NO 15
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Human
; US-09-402-532-15

Alignment Scores:
Pred. No.: 6,636-16 Length: 1146
Score: 220.50 Matches: 118
Percent Similarity: 34.90% Conservative: 60
Best Local Similarity: 23.14% Mismatches: 160
Query Match: 7.44% Indels: 173
DB: 3 Gaps: 27

US-09-651-651-5 (1-546) x US-09-402-532-15 (1-1146)
Qy      |||||
44  lIeIleIleProGLyPhEAlASeRThRgLUeUAlAArgLUlLeUSeRgLySeRPro 63
Db      |||||
31  GTGCTGGTCCCTGGTGTGATGGGTAAACCACTGGAAGCC-----AGCTGGACAAGCG 84
Qy      |||||
64  TyRThRProLeUSeRPhE-----ASeRProLeUSeRValTrPLUe 77
Db      |||||
85  ---ACAGTGTGCATCACTCTGCTGCTCAAGAAAGCAAGCAAGCTTACATTCAGCTG 141
Qy      |||||
78  ASeRThRlyLeUSeRAla---ValASeRgLyThRPhElySeRMeValLeUSeR 96
Db      |||||
142  AACCTGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 198
Qy      |||||
97  ProTyRAnGLnThRAnHlASeRProGLyLySeRgLySeAArgProASeR----- 112
Db      |||||
199  GTTCAACAACAAACATCC-----AGGCCAACCCAGTTCCGATGCTGTGAGTACGT 252
Qy      |||||
113  -----GlyLeUSeR-----AlAlleThRgLUeUaSeRProGLyTyRleThr--- 126
Db      |||||
253  GTCCCTGGCTTGGAGAAGCTTTCAGTGAAGTCTGAGACCCAGCAAAAGCAGCTG 312
Qy      |||||
127  Gly-ProLeUSeRThRValTrPLySeGLyThRPLySeRgLyTyRcyValGluPhEgLy---11 145
Db      |||||
313  GGTTCCTATTTCACACCATGCTG- GAGAAGCTT-----GTGGGTGGGCTACAC 362
Qy      |||||
145  eGLuAlASeNAIAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 165
Db      |||||
363  ACAGGAGTGAAGATGCCGAGGGCTCCCTATACCTGGGCGCGAGAGCCCA-----AA 413
Qy      |||||
165  uGLuAArgASeRlyUTrPhEAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 185
Db      |||||
414  TGAAGAGGGGCTTCTCTGCGGCTGCGGCTGCGGAGATGATCGAGAGATGTACAGCTG 473
Qy      |||||
185  gGLyGLyProSeRleIleAlASeRnHlASeRMeRgLySeAnValAPhEArgTyRPhE 205
Db      |||||
474  TGGGGCCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 533
Qy      |||||
205  uGUUUTPLUeAUGLeUgLUlLeAlAProLyshIeUlyeUTrPLUeApgLnHlE1 225
Db      |||||
534  GCAG-----CGGACGCCGACGGCTGGAAAGCAAGATATAT 569

```





PRIOR APPLICATION NUMBER: 60/080107	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194	PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334	PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796	PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336	PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742	PRIOR FILING DATE: 1998-04-29

1	PRIOR FILING DATE: 1998-04-30
2	PRIOR APPLICATION NUMBER: 60/084366
3	PRIOR FILING DATE: 1998-05-05
4	PRIOR APPLICATION NUMBER: 60/084414
5	PRIOR FILING DATE: 1998-05-06
6	PRIOR APPLICATION NUMBER: 60/084441
7	PRIOR FILING DATE: 1998-05-06
8	PRIOR APPLICATION NUMBER: 60/084637
9	PRIOR FILING DATE: 1998-05-07
10	PRIOR APPLICATION NUMBER: 60/084639
11	PRIOR FILING DATE: 1998-05-07
12	PRIOR APPLICATION NUMBER: 60/084660
13	PRIOR FILING DATE: 1998-05-07
14	PRIOR APPLICATION NUMBER: 60/084677
15	PRIOR FILING DATE: 1998-05-07
16	PRIOR APPLICATION NUMBER: 60/084663
17	PRIOR FILING DATE: 1998-05-07
18	PRIOR APPLICATION NUMBER: 60/085339
19	PRIOR FILING DATE: 1998-05-13
20	PRIOR APPLICATION NUMBER: 60/085338
21	PRIOR FILING DATE: 1998-05-13
22	PRIOR APPLICATION NUMBER: 60/085323
23	PRIOR FILING DATE: 1998-05-13
24	PRIOR APPLICATION NUMBER: 60/085582
25	PRIOR FILING DATE: 1998-05-15
26	PRIOR APPLICATION NUMBER: 60/085700
27	PRIOR FILING DATE: 1998-05-15
28	PRIOR APPLICATION NUMBER: 60/085689
29	PRIOR FILING DATE: 1998-05-15
30	PRIOR APPLICATION NUMBER: 60/085719
31	PRIOR FILING DATE: 1998-05-15
32	PRIOR APPLICATION NUMBER: 60/085580
33	PRIOR FILING DATE: 1998-05-15
34	PRIOR APPLICATION NUMBER: 60/085573
35	PRIOR FILING DATE: 1998-05-15
36	PRIOR APPLICATION NUMBER: 60/085704
37	PRIOR FILING DATE: 1998-05-15
38	PRIOR APPLICATION NUMBER: 60/085697
39	PRIOR FILING DATE: 1998-05-15
40	PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:	
Pred. No.:	2,83e-15
Score:	230.50
Percent Similarity:	34.90%
Best Local Similarity:	23.14%
Query Match:	73.44%
OB:	3
Length:	2680
Matches:	118
Conservative:	60
Mismatches:	160
Indels:	173
Gaps:	27

US-09-651-651-5 (1-546) x US-09-999-833A-156 (1-2680)

[illegible]



;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22

;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:  
Pred. No.: 2,83e-15  
Score: 220.50  
Percent Similarity: 34.90%  
Best Local Similarity: 23.14%  
Length: 2680  
Matches: 118  
Conservative: 60  
Mismatch: 160





Percent Similarity: 33.60%  
 Best Local Similarity: 22.60%  
 Query Match: 7.08%  
 DB: 3  
 Gaps: 25

US-09-651-651-5 (1-546) x US-09-489-847-57 (1-2687)

Qy ProGlyPheAlaSerThrGlnLeuAlaTrpSerIleLeuAspCysProTyrThrPro 66  
 165 CCAAGCTGACAGAC-----CGACAGTGGTCACTACTCTGCTCCAGAGACCG 215  
 Qy LeuAspPheAsnProLeuAspLeuValTyrLeuAspThrIlysLeuLeuSerAlaVal 86  
 216 AAGACTACTTCACAACTGGCTGAACCTG-----AAGTGTGCTGCTGTCGATCATT 269  
 Qy AsnCysTrpPheLeuAspCysMetValLeuAspProTyrAsnGlnThrAspHisProGluCys 106  
 270 GACTGCTGGATTGACAAATATCAGGCTG--GTTTACAAACAAACATCC-----AGGGCC 320  
 Qy LysSerArgProAspSer-----GlyLeuSer-----AlaIle 117  
 321 ACCACGTTTCTATATGCTGGATGATACGTGCTGCTGGCTTGGAGACCTTCTCACTG 380  
 Qy ThrGlnLeuAspProGlyTyrIleThr--Gly-ProLeuSerThrValTyrIysGluTr 136  
 381 GAGTTCCTGGACCCCAACAAAGCAGCGTGGTCTTATTTCCACACCATGCTG--GAGAG 439  
 Qy PheLeuTyrCysValGlnPheGly---IleGluAlaAsnAlaIleValAlaValProTyr 155  
 440 CCTT-----GTGGGCTGGGGCTACACACGGGGTGAAGATGTCGAGGGGCTCCTTA 490  
 Qy AspTrpArgLeuSerProThrIlysLeuGlnIuArgAspLeuTyrPheHisIlysLeuTyr 175  
 491 TGAATGGCGCGAGCCCA-----AATGAACAGGGGCTACTTCTCGGGCCCTCG 541  
 Qy PheLeuThrPheGlnThrAlaLeuIysLeuArgGlyGlyProSerIleValPheAlaHisSe 195  
 542 CGAGATGATCGAGAGATGATACGCTGTATGGGGGCGCCCTGTGCTGGTGGCCACAG 601  
 Qy rMetGlyAsnAsnValPheArgTyrPheLeuGlnTyrLeuArgLeuGlnIleAlaProTyr 215  
 602 TATGGGCAACATGTACACGCTTACTTGTGAG-----CG 637  
 Qy HisTyrLeuIysTrpLeuAspGlnHisIleHisAlaTyrPheAlaValAlaProLe 235  
 638 GCACCCGAGCGCCGGAARACAGATATCCGGGCTTCTGTCACAGCGGTGGCCCTG 697  
 Qy uLeuGlySerValGlnAlaIleIysSerThrLeuSerGlyValThrPheGlyLeuProVa 255  
 698 GGGGGGCGTGGCCAAAGCCCTGCGCTGCTGCTGAGGAGACAAACACCGATCCAGCT 757  
 Qy SerGlnGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeuTyrPheLeu 274  
 758 CATGGGGCCCTGAAGATCGGAGAGACACCGGCTGCTGTCTCCACAGCTGAGCTCT 817  
 Qy tProPheSerIysAsnCysIysGlyAspAsnThrSerTrpThrHisPheSerGlyGlyAl 294  
 818 GCCCTAC-----AACTACACATGCTCA----- 839  
 Qy AlaIysLeuAspIysArgValTyrHisCysAspGluGlnIuTyrGlnSerIysTyrSe 314  
 839 ----- 839  
 Qy rGlyTrpProThrAsnIleIleAsnIleGluIleProSerThrSerValThrGlnThrAl 334  
 840 -----CCTAGAGAGGCTTCTGTCGACAGACACC 865  
 Qy AlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaAr 354  
 866 CACATCAACTACACA----- 881  
 Qy gGluLeuAlaAspGlyThrIleuPheLeuAlaIleGluIysPryTyrAspProAspSerIysAr 374  
 354 ----- 881

Db 882 -----CTGGGGGACTAC----- 893  
 Qy 374 gMetLeuHisGlnLeuIysLeuTyrHisAsp----- 385  
 Db 894 -----CGCAGAGTCTTCCAGACATCGGCTTGAAGATGGCTGGCTCAT 937  
 Qy 386 -----AspProValPheAsnProLeuThrProTyrGluIuArgProProIleIysAsnVa 403  
 Db 938 GCGGAGACACACA-----GAAAGGCTGTGGAAGCCAGATGCCACTGGCTGCAGCT 991  
 Qy 403 lPheCysIleTyrGlyAlaHisIleuIysThrGlnValGlyTyrTyrPheAlaProSerG 423  
 Db 992 GCACCTGCTCTATGCT-----ACTGG 1012  
 Qy 423 YLysProTyrProAspAsnTrpIleIleThrAspIleIleTyrGlnThrGlnIysSerLe 443  
 Db 1013 CGTCCCCACACCGACTCTCTC-----TACTATGAG-----AGCTT 1048  
 Qy 443 uValSerArgSerGlyThrValValAspGlyAsnAlaGlyProIleThrGlyAspGluThr 463  
 Db 1049 CCTGACCTGACCTTAATAATCTGCTTGGTGC-----GCCATGTGAC 1093  
 Qy 463 rValProTyrHisSerLeuSerTrpCysIysAsnTrpLeuGlyProIysValAsnIleThr 483  
 Db 1094 TGTGAACCTGAAGATGCCCTGCAGTGCAGGCTGGCAGAGCCGACAGAGACCAAGT 1153  
 Qy 483 rMetAlaProGlnProGluHisAspGlySerAspValHisValGluLeuAsnValAspHis 503  
 Db 1154 GTTGCTG-----CAGGAGCTGCCAGCAGCAG--CACATGAGATG----- 1193  
 Qy 503 sGluHisGlySerAspIleIleAlaAsnMetThrIysAlaProArgValIysTyrIle 522  
 Db 1194 -----CTGGCCAAAGCCACACCTGGCTTACTGTGAACGTGTG 1232

RESULT 13  
 US-09-402-532-13  
 ; Sequence 13, Application US/09402532  
 ; Patent No. 6498019  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Taniyama, Yoshio  
 ; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE  
 ; FILE REFERENCE: 2456USOP  
 ; CURRENT APPLICATION NUMBER: US/09/402,532  
 ; EARLIER FILING DATE: 1999-10-04  
 ; EARLIER APPLICATION NUMBER: PCT/JP98/01643  
 ; EARLIER FILING DATE: 1998-04-09  
 ; EARLIER APPLICATION NUMBER: JP 10-010289  
 ; EARLIER FILING DATE: 1998-01-22  
 ; EARLIER APPLICATION NUMBER: JP 9-184885  
 ; EARLIER FILING DATE: 1997-07-10  
 ; EARLIER APPLICATION NUMBER: JP 9-093355  
 ; EARLIER FILING DATE: 1997-04-11  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE:  
 ; SEQ ID NO 13  
 ; LENGTH: 1233  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-402-532-13

Alignment Scores:  
 Pred. No.: 1,78e-14  
 Score: 209.00  
 Percent Similarity: 33.79%  
 Best Local Similarity: 22.46%  
 Query Match: 7.05%  
 DB: 3  
 Gaps: 25

US-09-651-651-5 (1-546) x US-09-402-532-13 (1-1233)

Qy 39 SerIysLeuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSer 58  
 158 TCATTGACTGCTGATTTGACAAATATCAGGCTGAGTGCAGTG-----GCCCAATCT 208



```

Db      272 -----CGATAACTGGGATTAACAGCGTGGTT----- 298
Qy      97 ProTyraSnglnThrAspHisProGluCysLysSer----- 108
Db      299 -----ACAACAAACATCCAGAGGCCACCCAGTTTCTGTATGGTGGAATGACGTC 349
Qy      109 -----ATGProAspSerGlyLeuSerAlaIleThrGluLeuAspProGly 123
Db      350 TCCCTGGCTTTGGGAAGACCT---TCTCACTGAGATTCCCGACCCCGAAGAAAGACGCG 406
Qy      124 TyrIleThrGlyProLeuSerThrValTrpLysGluTrpLeuLysTrpCysValGluPhe 143
Db      407 TGGGTT-----CCTATTTCACACCATGGTG--GAGAGCCTT-----GTGGGCTGG 450
Qy      144 Gly---IleGluAlaSerAlaIleValAlaValProTyraSnpTrpArgLeuSerProThr 162
Db      451 GGGTACACACGGGGGTGAGATGTCGAGGGGCTCCATGACTGGCGCGAGCCCA--- 507
Qy      163 LysLeuGluGluTrpArgLeuThrPheHisLysLeuLysLeuThrPheGluTrpAlaLeu 182
Db      508 -----AATGAAACGGGCGCTTACTTCTGGCCCTCGCGAGATGANTCGAGAGATGTAC 561
Qy      183 LysLeuArgGlyGlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArg 202
Db      562 CAGCTGATGGGGGCCCGCTGGTGCTGTGGCCACAGTATGGGCAACATGTACACGCTC 621
Qy      203 TyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisLysLysLysTrpLeuAsp 222
Db      622 TACTTCTGCAG-----CGGACGCCGCGAGCCTCGAAGGAC 657
Qy      223 GlnHisIleHisIleAlaTyrPheAlaValAlaProLeuLeuGlySerValGluAlaIle 242
Db      658 AAGATATCCGGGCTTCGTGTCTACTGTGGCGCCCTGGGGGGCGCGGCGCAAGACCTG 717
Qy      243 LysSerThrLeuSerGlyValThrPheGlyLeuProVal---SerGluGlyThrAlaArg 261
Db      718 CGGCTCTGGCTTCAGAGACACACACCGGATCCCAATGATCCGGGCCCTGAAGATCCGG 777
Qy      262 LeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCysLys 281
Db      778 GAGCAGACGGGCTAGCTGTCTCCACAGCTGGCTGCTGCTAC----- 822
Qy      282 GlyAspAsnThrSerTrpThrHisPheSerGlyAlaAlaLysLysAspLysArgVal 301
Db      823 -----AATACACATGCTCA----- 837
Qy      302 TyrHisCysAspArgGluGluTrpGlnSerLysTyrSerGlyTrpProThrAsnIleIle 321
Db      837 ----- 837
Qy      322 AsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsnMetThrSerMet 341
Db      838 -----CCTGAGAAAGCTGTCTGTGCAGACACCCACACATCACTACACA----- 879
Qy      342 GluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeu 361
Db      879 ----- 879
Qy      362 PheLysAlaIleGluAspTrpTyrAspProAspSerLysArgMetLeuHisGlnLeuLysLys 381
Db      880 -----CTCGCGGAGCTAC-----CGCAG 897
Qy      382 LeuTyrHisAsp-----AapProValPheAsn 390
Db      898 TTCTTCCAGACATCGGCTTGAAGATGGCTGGCTCATGGCGACAGACACA-----GAA 951
Qy      391 ProLeuThrProTrpGluArgProProIleLysAsnValPheCysIleTyrGlyAlaHis 410
Db      952 GGGCTGGTGAAGCAGATGCACTGCGTGGACGCTGCACTGCTTATGGT----- 1005
Qy      411 LeuLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAsnTrp 430

```

```

Db      1006 -----ACTGGGTCCCAACACAGACTCTTC 1032
Qy      431 IleIleThrAspIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyThrVal 450
Db      1033 -----TACTATGAG-----AGTTCCTCGACCGTGAACCTTAAATC 1068
Qy      451 ValAspGlyAsnAlaGlyProIleThrGlyAspArgLysValProTyrHisSerLeuSer 470
Db      1069 TCGTTTGGTGAC-----GGCGATGGTACTGTGAACCTTGAAGAGTCCCTG 1113
Qy      471 TrpCysLysAsnTrpLeuGlyProLysValAsnIleThrMetAlaProGlnProGluHis 490
Db      1114 CAGTGGCAGCGCTGGAGAGCGCCGACAGACACCAAGTTTCTG-----CAGGAGCTG 1167
Qy      491 AspGlySerAspValHisValGluLeuAsnValAspHisGlnHisGlySerAspIleIle 510
Db      1168 CCAGCGAGGAG--CACATCGAGATG-----CTG 1194
Qy      511 AlaAsnMetThrLysAlaProArgValLysTyrIle 522
Db      1195 GCCACGCCACACACCTGGCTTATGAAACGTGTG 1230

RESULT 15
US-09-402-532-18
; Sequence 18, Application US/09402532
; Patent No. 6498019
; GENERAL INFORMATION:
; APPLICANT: Taniyama, Yoshio
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
; FILE REFERENCE: 2456USOP
; CURRENT APPLICATION NUMBER: US/09/402, 532
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/JP98/01643
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: JP 10-010289
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: JP 9-184885
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 9-093355
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE:
; SEQ ID NO 18
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Human
US-09-402-532-18

Alignment Scores:
Pred. No.: 2, 03e-14 Length: 1332
Score: 209.00 Matches: 115
Percent Similarity: 33.79% Conservative: 58
Best Local Similarity: 22.46% Mismatches: 154
Query Match: 7.05% Indels: 186
DB: 3 Gaps: 25

US-09-651-651-5 (1-546) x US-09-402-532-18 (1-1332)
Qy      39 SerLysLeuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSer 58
Db      257 TCAATTGACTGCTGATTTGACAAATATCAGCTGGAGTGCAGT-----GGCAATCT 307
Qy      59 IleLeuAspCysAspProTyrThrPro-----LeuAspPheAsnProLeuAspLeuValTrp 76
Db      308 CGGCTCACTACACTGCTGCTCCAGGCTCAAGACACTTCTCGCTGCTACAGCTC----- 361
Qy      77 LeuAspThrThrLysLeuLeuSerAlaValAsnCysTrpPheLysCysMetValLeuAsp 96
Db      362 -----CGATACTGGGATTAACAGCGTGGTT----- 388
Qy      97 ProTyraSnglnThrAspHisProGluCysLysSer----- 108
Db      389 -----ACAACAAACATCCAGAGGCCACCCAGTTTCTGTATGGTGGAATGACGTC 439

```

```

QY 109 -----ArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGly 123
Db 440 TCCCTGGCTTTGGGAAGACCT---TCTCAGTGAAGTTCTGGAGCCCAAGACAGCG 496
QY 124 TyrIleThrGlyProLeuSerThrValTrrlySGluTrpLeuIleTrpCysValGluPhe 143
Db 497 TGGGTT-----CCTATTTCACACCATGGTG-GAAGACCTT-----GTGGGCTGG 540
QY 144 Gly---IleGluAlaAsnAlaIleValAlaValProTyrAspTrpArgLeuSerProThr 162
Db 541 GCGTACACACGGGGGTGAGGATGTCGAGGGGCTCCCTATGACTGGGCGCGACCCCA--- 597
QY 163 LysLeuGluGluArgPheLeuTyrPheHisLysLeuLysLeuThrPheGluThrAlaLeu 182
Db 598 -----AATGAAACCGGGCCCTAATTCTGGCCCTCCGGAGATGATTCAGAGATGTAC 651
QY 183 LysLeuArgGlyGlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArg 202
Db 652 CAGCTGTATGGGGGCGCCCTGGTGGTGGTCCCAAGATGGGCAACATGTACACGCTC 711
QY 203 TyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHisTyrLeuLysTrpLeuAsp 222
Db 712 TACTTTCTGCAG-----CGCAGCGCGCAGCGCTGGAAGGAC 747
QY 223 GlnHisIleHisAlaIleTyrPheAlaValAlaProLeuLeuGlySerValGluAlaIle 242
Db 748 AAGTATATCGGGCTTCAGAGACAACAACCGGATCCAGTCACTGGGGGCGCTGGCAAGATCGGG 807
QY 243 LysSerThrLeuSerGlyValThrPheGlyLeuProVal---SerGluGlyThrAlaArg 261
Db 808 CCGCTCTCGGCTTCAGAGACAACAACCGGATCCAGTCACTGGGGGCGCTGGCAAGATCGGG 867
QY 262 LeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnCysLys 281
Db 868 GAGCAGCAGCGGTCAAGCTGCTCCACACCACTGGCTGCTGCCCTAC----- 912
QY 282 GlyAspAsnThrSerThrPheHisPheSerGlyAlaAlaLysLysAspLysArgVal 301
Db 913 -----AACTACACATGTCA----- 927
QY 302 TyrHisCysAspGluGluTyrGlnSerLysTyrSerGlyTrrProThrAsnIleIle 321
Db 927 ----- 927
QY 322 AsnIleGluIleProSerThrSerValThrGluThrAlaLeuValAsnMetThrSerMet 341
Db 928 -----CCTAGAAAGTGTTCGTGCAGACACCAACATCAACTACACA----- 969
QY 342 GluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeu 361
Db 969 ----- 969
QY 362 PheLysAlaIleGluAspTyrAspProAspSerLysArgMetLeuHisGlnLeuLysLys 381
Db 970 -----CTGGGGGACTAC-----CGCAAG 987
QY 382 LeuTyrHisAsp-----AspProValPheAsn 390
Db 988 TTCTTCAGAGACATCGGCTTTGAAGATGGCTGCTCATGGCGGAGAGACACA-----GAA 1041
QY 391 ProLeuThrProTrpGluArgProProIleLysAsnValPheCysIleTyrGluAlaHis 410
Db 1042 GGGCTGTGAGAACCCACGATGCGACCTGGGCTGACGTGCTGCTTATGGT----- 1095
QY 411 LeuLysThrGluValGlyTyrTyrPheAlaProSerGlyLysAspProTyrProAspAsnTrp 430
Db 1096 -----ACTGGGCTCCCAACACAGACTCCTTC 1122
QY 431 IleIleThrAspIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyThrVal 450
Db 1123 -----TACTATAG-----AGCTTCCTGACCGTGAACCTTAAATC 1158

```

```

QY 451 ValAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSerLeuSer 470
Db 1159 TGCCTTGGTGAC-----GCGATGGTACTGTGAACCTTGAAGATGCCCTG 1203
QY 471 TrpCysLysAsnTrpLeuGlyProLysValAsnIleThrMetAlaProGlnProGluHis 490
Db 1204 CAGTCCGAGGCTGTGGCAGAGCCGCCAGAGACACCAAGTGTGCTG-----CAGGAGCTG 1257
QY 491 AspGlySerAspValHisValGluLeuAsnValAspHisGlnHisGlySerAspIleIle 510
Db 1258 CCAAGCAGCGAG---CACATCGAAGTG-----CTG 1284
QY 511 AlaAsnMetThrLysAlaProArgValLysTyrIle 522
Db 1285 GCCAAGCCACACACCTGGCTATCTGAACGTGTG 1320

```

Search completed: November 22, 2005, 13:54:28  
Job time : 237 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: November 22, 2005, 10:59:19 ; Search time 834 Seconds  
(without alignments)  
5413.765 Million cell updates/sec

Title: US-09-651-651-5  
Perfect score: 2965  
Sequence: 1 MGANSKSYATSAFTVIAVFL.....DSEISPGKRTAVWEIDKSGY 546

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-DB=published Applications\_NA\_Main -GPM=faetap -SUFFIX=rnpbm -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62  
-TRANS-human40.caf -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09651651 @CGN 1 1 797 @rnat 18112005 101832\_20311  
-NCPU=6 -ICPU=3 -NO MAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCk=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRESH=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELXT=6 -DELXT=7

Database : Published Applications\_NA\_Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1953	65.9	1719	7	US-10-437-963-76086
2	595	20.1	1760	7	US-10-424-599-16008
3	569.5	19.2	528	7	US-10-424-599-134518
4	482	16.3	848	8	US-10-425-115-93907
5	482	16.3	924	8	US-10-425-115-137717
6	461	15.5	1383	8	US-10-425-115-93909
7	325.5	11.0	2326	8	US-10-883-760-45
8	309	10.4	1433	7	US-10-424-599-17331

9	284	9.6	2398	6	US-10-321-802-25	Sequence 25, Appl
10	283.5	9.6	2700	6	US-10-321-802-23	Sequence 23, Appl
11	281	9.5	2030	6	US-10-321-802-27	Sequence 27, Appl
12	280	9.4	2479	6	US-10-321-802-17	Sequence 17, Appl
13	274	9.2	2395	7	US-10-437-963-38278	Sequence 38278, A
14	272.5	9.2	2700	8	US-10-425-115-55049	Sequence 55049, A
15	269	9.1	4093	6	US-10-321-802-35	Sequence 35, Appl
16	261.5	8.8	2433	6	US-10-321-802-21	Sequence 21, Appl
17	257.5	8.7	2656	6	US-10-321-802-19	Sequence 19, Appl
18	255.5	8.6	2425	9	US-10-509-796-1	Sequence 1, Appl
19	253	8.5	2838	7	US-10-424-599-125713	Sequence 125713, A
20	234	7.9	1137	6	US-10-323-051-14	Sequence 14, Appl
21	234	7.9	1236	6	US-10-323-051-19	Sequence 19, Appl
22	234	7.9	1236	7	US-10-478-466-2	Sequence 2, Appl
23	232	7.8	1998	6	US-10-217-939-5	Sequence 5, Appl
24	226	7.6	1744	3	US-09-919-497-30	Sequence 30, Appl
25	226	7.6	1744	3	US-09-880-107-2245	Sequence 2245, Ap
26	220.5	7.4	1137	6	US-10-323-051-12	Sequence 12, Appl
27	220.5	7.4	1146	6	US-10-323-051-15	Sequence 15, Appl
28	220.5	7.4	1236	6	US-10-323-051-17	Sequence 17, Appl
29	220.5	7.4	1400	7	US-10-755-889-577	Sequence 577, App
30	220.5	7.4	2680	3	US-09-978-295A-156	Sequence 156, App
31	220.5	7.4	2680	3	US-09-978-697-156	Sequence 156, App
32	220.5	7.4	2680	3	US-09-978-189-156	Sequence 156, App
33	220.5	7.4	2680	3	US-09-978-608A-156	Sequence 156, App
34	220.5	7.4	2680	3	US-09-978-585A-156	Sequence 156, App
35	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
36	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
37	220.5	7.4	2680	3	US-09-978-403A-156	Sequence 156, App
38	220.5	7.4	2680	3	US-09-978-564A-156	Sequence 156, App
39	220.5	7.4	2680	3	US-09-978-833A-156	Sequence 156, App
40	220.5	7.4	2680	3	US-09-981-915A-156	Sequence 156, App
41	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
42	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
43	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
44	220.5	7.4	2680	3	US-09-978-824-156	Sequence 156, App
45	220.5	7.4	2680	3	US-09-978-423A-156	Sequence 156, App

## ALIGNMENTS

RESULT 1  
US-10-437-963-76086  
; Sequence 76086, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Bouharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221) B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 76086  
; LENGTH: 1719  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76111C.1  
US-10-437-963-76086  
Alignment Scores:  
Pred. No.: 5.41e-226  
Score: 1953.00  
Percent Similarity: 74.02%  
Best Local Similarity: 65.30%  
Length: 1719  
Matches: 367  
Conservative: 49  
Mismatch: 82

Query Match:	65.87%	Indels:	64
DB:	7	Gaps:	7

US-09-651-651-5 (1-546) x US-10-437-963-76086 (1-1719)

```

QY      7 SerValTThAlAserPheThrValIleAlaValPhePheLeuIleCyGlyGlyArgThr 26
         |||||
Db      40 TCGATGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGC 99
         |||||

QY      27 AlaValGIuaBPduThrGIuPhenIGlyAspTYrSerLyLeuSergIyIleIle 46
         ::::
Db      100 GGAGGCGATGGCGGCGCGGATTTC-----GACTACCGGAAGCTCTCGGGATTAATATC 153
          |::|

QY      47 ProGlyPheAlaSerThrGIuLeuAlaITPrSerILEuLusPCyProTYrThrPro 66
         |||||
Db      154 CCGGGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCC 213
          |::|

QY      67 LeuAspPheAsnProLeuAspLeuValITPrLeuAspThrThryLyLeuSeraIAlaI 86
         |||||
Db      214 TTGCAGTTAAACCCCTCGATCTCGTGCTGGCTGCACGCCCAAGCTCTTATCTGCTGC 273
          |::|

QY      87 AsnCysTrpPheLYsCysMetValLeuAspProTYrAsnGlnThrAspHisProGluCys 106
         |||||
Db      274 AATTGGTGGCTCAATATGATGATGCTGCTGCACCTTACAACACGAGCATCCAGAGTGC 333
          |::|

QY     107 LySserArgProAspSergIyLeuSeraIleThrGIuLeuAspProGlyTYrIleThr 126
         |||||
Db     334 AAGTCAAAGCCCTGACAGTGGTCTTTGTGCATAATACAGAGTTAACCTGGTGTATATAACT 393
          |::|

QY     127 GlyProLeuSerThrValITrplySGuITrplySGuITrplyStrycVaValGIuPheGlyIleGlu 146
         |||||
Db     394 GGTCTCTTCTTCTCACTGGGAAAGAATGGGTCAAATGG----- 432
          |::|

QY     147 AlaAsnAlaIleValAlaValProTYrAspTrpArgLeuSerProThryLyLeuGIuGlu 166
         |||||
Db     432 ----- 432

QY     167 ArgAspLeuTyrrHeHibLySbleuLybleuThrPheGIuThrAlaLeuLySleuArgGly 186
         |||||
Db     433 -----TTAAGCTTTGAAACTGATTTGAACCTTCAGNGA 465
          |::|

QY     187 GIYProSerILEvalPheAlaHisSerMetGLYAsnAsnValPheArgTYrPheLeuGIu 206
         |||||
Db     466 GGGCCTCTTATGATGTTGGTCTCATTCAGTAGGTAATATGTTGGTCACTTTCGGAA 525
          |::|

QY     207 TrpLeuArgLeuGIuIleAlaPdybHSYrLeuLySTrPLeuAspGINHisIleHis 226
         |||||
Db     526 TGGTGTAAACTGAAATAATCGCTCCCAAGCATTAATCGAGTGGCTGACGAACATATACAT 585
          |::|

QY     227 AlaTYrPheAlaValaGIuAProLeuLeuGlySerValGIuAlaIleLySserThrLeu 246
         |||||
Db     586 GCATTACTTGGCACTGGTGCACTCTTCTTGTTCTTACTGAAACAGTTAAAGCTGCTTT 645
          |::|

QY     247 SergIyValIThrPheGlyLeuProValSergIyGlyThrAlaArgLeuSeraSnSer 266
         |||||
Db     646 TCTGGAGCAACATTTGGTCTTCCAGTCAGCGAAGGAACACACAGATTGAATTAATGCA 705
          |::|

QY     267 PheAlaSerSerLeuITrPLeuMetProPheSerLyLeuAsnCylAspArgAnthrSer 286
         |||||
Db     706 TTGTGTCAATCTTTAATGAGCTCATGCCCTTCTCGAATATTTGCAAAAGCTGTATATATAC 765
          |::|

QY     287 TrpThriSPhe-----SergIyGlyAlaIleLybVAspLybSAvgValTYrHis 303
         ::::
Db     766 TGGAGCATTTCTTTGAGGGAAGGAGGTGGCCACACACACAGCA----- 813
          |::|

QY     304 CyAspGIuGIuGluITyrGInserLySYrSergIYTPrProThrAsnIleIleAsnIle 323
         |||||
Db     814 TGTGATGAATATGGAATATATAGTCTGAATACTACAGAGTGGCCCAAACTTTCAGCATC 873
          |::|

QY     324 GluIleProSerThrSeraValIThrGIuThr-----AlaLeuValAsn 337
         |||||
Db     874 GAGGTCTTCAAGATTCGAGTATAGGAAAGCATACCATCATTAATGATACACTGAAATC 933
          |::|

```

[illegible]

Score: 595.00 Matches: 105  
Percent Similarity: 87.77% Conservative: 17  
Best Local Similarity: 75.54% Mismatches: 15  
Query Match: 20.07% Indels: 2  
DB: 7 Gaps: 1

US-09-651-651-5 (1-546) x US-10-424-599-16008 (1-1760)

```
Oy 410 Hleleu-----lyerhrgluvalgllytyrphenalaprosergilylvsprotyrpro 427
Db 71 CATCTTGACTTTAAAAAATAGCTTGTTACTACTATAGCTCTAGCGGCAAGCTTACCC 130
Oy 428 AspAsnTrpIleIleThraSpIleIleTyrgluThrgluGlySerleuValSerAgSer 447
Db 131 GATACATCGATCATTAACGATGTCGTTATGAGTTGAAAGATCTTAATCTCAAGGTCA 190
Oy 448 GltThraValValAspGlyAsnAlaGlyProIleThrgluAspGluThraValProtyrHis 467
Db 191 GGGAAATCTGGTTGAAGGAAACCTGAGCAATTAAGTCGATGAGCGGTGCCATATCTC 250
Oy 468 SerIleuSerTrpCysIleAsnTrpLeuGlyProIleValAsnIleThraMetAlaProGln 487
Db 251 TCCCTTTCTCGTGCCAAAACTGGCTTGGACCCAAAGTGAACATTAACAAAAGCTCCACAG 310
Oy 488 ProGluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySer 507
Db 311 TCAGAGCATGATGGCTCAGATGTCAATTAATTAATTTGGAACATCAACATGAAGAA 370
Oy 508 AspIleIleAlaAsnMetThrValAlaProtyrValIleThraPheTyrgluAsp 527
Db 371 GATATCTTCCCAACATGACAGCAAGATCAACAAAGGTGAAGTATTAACATTTATGAAGAT 430
Oy 528 SerGluSerIleProGlyIleTyrgluThraValIleThrgluLeuAspGlySerGlyTyr 546
Db 431 TCTGAAGCTCTTCTCGAAGAAAGAGACAGCACTTTGGAGCTTGATTAAGCAATCTAC 487
```

RESULT 3

US-10-424-599-134518  
; Sequence 134518, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 134518  
; LENGTH: 528  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_9247C.1  
US-10-424-599-134518

Alignment Scores:

Pred. No.: 1,39e-58 Length: 528  
Score: 569.50 Matches: 105  
Percent Similarity: 87.60% Conservative: 8  
Best Local Similarity: 81.40% Mismatches: 12  
Query Match: 19.21% Indels: 4  
DB: 7 Gaps: 2

US-09-651-651-5 (1-546) x US-10-424-599-134518 (1-528)

```
Oy 17 ValPhePheIleuIleCysGlyGlyAlaXrThraValAlaGluAspGluThrgluPheHisGly 36
Db 152 GTCTCTCTGCGCGGTGCTGCGGAGCCCACTCA---GACGACGGGCGGAGCTC----- 202
```

Oy 37 AspTySerIleuSerGlyIleIleIleProGlyIlePheAlaSerThrgluLeuArgAla 56  
Db 203 GACTATCCAAACCTCTCCGCGCATTTAATATCCCGGCTTGCGATCCACTGCGAGCA 262

Oy 57 TrpSerIleLeuAspCysProIleThraProLeuAspPheAsnProLeuAspLeuValTrp 76  
Db 263 TGGTCAATCTCGACGCTGCTTACTCTCCGCTCGATTTCAACCTCTGACCTTGCGTGG 322

Oy 77 LeuAspThrThrgluLeuSerAlaValAlaAsnCysTrpPheIleCysMetValLeuAsp 96  
Db 323 CTCGACACAAACAAACTTCTTCTGCTGTCATATGCTGCTTAAGTGAAGTGTGTGGAT 382

Oy 97 ProTyAsnGluThraAspHisProGluCysIleSerArgProAspSerGlyLeuSerAla 116  
Db 383 CCTTACAAACCAACAATCATCATCTGATTCAGAGTCCCGTCCGATAGTGCTTCTTGGC 442

Oy 117 IleThGluLeuAspProGlyTyrlleThrgluProIleSer-ThraValTrpIleGluTrp 136  
Db 443 ATTACGAACCTTGATTCAGGTTATTAACAAGACCTTTCAATTCAGTTGGAAAGAAAG 502

Oy 136 PleuTyTrpCysValGluPheGly 144  
Db 503 GATTAAGTGTGTGATTGAATTGGC 527

RESULT 4

US-10-425-115-93907  
; Sequence 93907, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 93907  
; LENGTH: 848  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_185630C.1  
US-10-425-115-93907

Alignment Scores:

Pred. No.: 1.3e-47 Length: 848  
Score: 482.00 Matches: 85  
Percent Similarity: 88.68% Conservative: 9  
Best Local Similarity: 80.19% Mismatches: 8  
Query Match: 16.26% Indels: 4  
DB: 8 Gaps: 1

US-09-651-651-5 (1-546) x US-10-425-115-93907 (1-848)

```
Oy 69 PheAsnProLeuAspLeuValTrpLeuAspThrThrgluLeuSerAlaValAsnCys 88
Db 358 TTCACTTTCTTACGCTCATTTACAT-----GCTTCTCTGCGGTAATATGCG 405
```

Oy 89 TrpPheIleCysMetValLeuAspProIleTyrgluThraAspHisProGluCysIleSer 108  
Db 406 TGGCTTAATGATGCTGCTTGAACCTTATATCAGATGACCATCCCGAATGCCAAGTCA 465

Oy 109 ArgProAspSerGlyLeuSerAlaIleThrgluLeuAspProGlyTyrlleThrgluPro 128  
Db 466 AGGCTGATGCTGCTTCTGCAATTAACAGCTGAGCTGCTGTTATTAACAGGTCT 525

Oy 129 LeuSerThraValTrpIleGluTrpLeuIleTyrgluValGluPheGlyIleGluAlaAsn 148  
Db 526 CTCTCTTCATTAAGAAAGAAATGGGTCAATATGGTGTAGTTGGCATTTGAAGCTAAT 585



```
QY      149 AlaIleValAlaValProTyrAserPheLeuSerProThrIlyLeuGluIuArgAap 168
      |||.....|
Db      586 GCAATTATCGCTGTTCCGTAATGAGACTGCCCATCAATGCTTGAGAGAGAT 645
QY      169 LeuTyrPheHisIlyLeu 174
      |||.....|
Db      646 CTGACTTTCACAAATTA 663

RESULT 5
US-10-425-115-137717
; Sequence 137717, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 137717
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57075C.1
US-10-425-115-137717

Alignment Scores:
Pred. No.:      1,496-47      Length:      924
Score:          482.00      Matches:      83
Percent Similarity: 94.68%      Conservative: 6
Best Local Similarity: 88.30%      Mismatches: 5
Query Match:    16.26%      Indels:      0
DB:             8           Gaps:      0

US-09-651-651-5 (1-546) x US-10-425-115-137717 (1-924)
QY      82 LeuLeuSerAlaValAsnCysTyrPheIlySyrMetValLeuAspProTyrAenGlnThr 101
      |||.....|
Db      595 CTTTCTCTGCGGTAATGCTGCTTAATGATGCTGTTAAACCTATATATACAGATA 654
QY      102 AspHisProGluCysIlySerArgProAspSerGlyLeuSerAlaIleThrGluLeuAap 121
      |||.....|
Db      655 GACCATCCGGAATGCAAGCTCAAGCCTGATAGTGTCTTCTGCAATTACAGAGCTGAGAC 714
QY      122 ProGlyTyrIleThrGlyProLeuSerThrValITrIlyLeuGluTrIlyLeuSyrProCysVal 141
      |||.....|
Db      715 CTGGTTATATACAGGCTCTCTCTCTTCAAGTATGGAAGAAAGGCTCAAAATGCTGTGA 774
QY      142 GluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAserPheLeuSerPro 161
      |||.....|
Db      775 GAGTTTGGAGTGAAGTATGCAATGATGCTGTTCCGATGATGAGAGACTGCCCA 834
QY      162 ThrIlyLeuGluIuArgAserIlyTyrPheHisIlyLeuIlyS 175
      |||.....|
Db      835 TCAATGCTTGAGAGAGATCTGTACTTTCACAAATTAAG 876

RESULT 6
US-10-425-115-93909
; Sequence 93909, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 93909
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_185632C.1
US-10-425-115-93909

Alignment Scores:
Pred. No.:      1,026-44      Length:      1383
Score:          461.00      Matches:      110
Percent Similarity: 50.19%      Conservative: 19
Best Local Similarity: 42.80%      Mismatches: 43
Query Match:    15.55%      Indels:      85
DB:             8           Gaps:      8

US-09-651-651-5 (1-546) x US-10-425-115-93909 (1-1383)
QY      59 IleLeuAspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValITrLeuAsp 78
      |||.....|
Db      660 TTGTTGGAATGCTCTTCAACTTTCCTTA-----CGCTCG 692
QY      79 ThrTrIlyLeuLeuSerAlaValAsnCysTyrPheIlySyrMetValLeuAspProTyr 98
      |||.....|
Db      693 TTTACATGCTTTTCTCTACGGTAATGCTGCTTAATGATGATGCTGTTGAACCTAT 752
QY      99 AenGlnThrAspHisProGluCysIlySerArgProAspSerGlyLeuSerAlaIleThr 118
      |||.....|
Db      753 AATCAGACAGACCATCCGGAATGCAATCAAGCCTGATAGTGTCTTTCGCAATTACA 812
QY      119 GluLeuAspProGlyTyrIleThr-----GlyProLeuSerThrValITrIlyGluTr 136
      |||.....|
Db      813 GAGCTGAGACCCGCTTAATATACAGGTTCAAGTCTCTCTTCAAGTATGGAAGAAATG 872
QY      136 PheLeuSyrProCysValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAs 156
      |||.....|
Db      873 GGTCAAAATGCTGTACAGTTTGGCATTTGAACCTATATACGATTCGCTGTTCCGTATGA 932
QY      156 PTrIlyLeuSerProThrIlyLeuGluIuArgAserIlyTyrPheHisIlyLeuIlyLe 176
      |||.....|
Db      933 TTGAGAGTCCGCCCATCATGCTTGAAGAGAGATCTGTACTTTCACAAATTAAGTT 992
QY      176 U----- 176
Db      993 TGTAACTTGCCTCACTTGTATGAACCAACCAATGCTATACATCTTTAGGATCAG 1052
QY      177 -----ThrPheGluThrAl 181
      |||.....|
Db      1053 TAAAGTTAATGCCCAAGTACAGGATTCAGATATCCTGCTACCAACAGATCCCAAG 1112
QY      181 AleuIlyLeuArgIlyPro----- 188
      |||.....|
Db      1113 CATACGGTTACCGCAATGCTGTGCAAGTTGACAGTACCAACCATCAGAGCCATAG 1172
QY      189 -----SerIleValPheAlaHisSe 195
      |||.....|
Db      1173 GAGGTTACACGCTGAGTACAGAGAAACATTAATTGAAAGATGGCAATTTGAAGAG 1232
QY      195 rMetGlyAsnAsnValPheArgTyrPheLeuGluTrIlyLeuGluIuIleAlaProIly 215
      |||.....|
Db      1233 TTACTCGAATCCAAATTTTGTCTTTT---GAATTCCTGGTGTACACCTGCAAAATAT 1289
QY      215 HisIlyTyrLeuIlyS-----TrIlyLeuAspGlnHisIleHisAlaIlyTyrPheAlaValGlyAl 233
      |||.....|
Db      1290 AACCATATTGAAGAGATGTCATTTGCATCATCATCTCATTTGTACTTT----- 1338
QY      233 aProLeuLeuGlySerVal-----GluAlaIleIlySerThrIleu 246
      |||.....|
Db      1339 -----TTGGGTGAGATCTAATAAAGCATGCAATTAAATTTACAAATG 1380
```

RESULT 7  
US-10-883-760-45  
; Sequence 45, Application US/10883760  
; Publication No. US2005035174A1  
; GENERAL INFORMATION:  
; APPLICANT: Yadav, Narendra  
; APPLICANT: Zhang, Hongxiang  
; TITLE OF INVENTION: ACYLTRANSFERASES FOR ALTERATION OF POLYUNSATURATED FATTY ACIDS  
; FILE REFERENCE: C12302 US NA  
; CURRENT APPLICATION NUMBER: US/10/883,760  
; CURRENT FILING DATE: 2004-07-06  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 45  
; LENGTH: 2326  
; TYPE: DNA  
; ORGANISM: Yarrowia lipolytica  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2271)..(2271)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-883-760-45  
  
Alignment Scores:  
Pred. No.: 7,01e-28 Length: 2326  
Score: 325.50 Matches: 119  
Percent Similarity: 38.14% Conservative: 66  
Best Local Similarity: 24.54% Mismatches: 202  
Query Match: 10.98% Indels: 99  
DB: 8 Gaps: 13  
  
US-09-651-651-5 (1-546) x US-10-883-760-45 (1-2326)  
  
Qy 27 AlaValGluuSerGluThrGluPheHisGlyAspTyrSerLysLeuSerGlyIleIle 46  
Db 709 GCTGTGGCGAAGCTATGAAAGCGAGAGACTCAACCCGAAGTACCGCGTGGTCTGTG 768  
Qy 47 ProGlyPheAlaSerThrGlnLeuArgAlaTyrSerIle-----LeuAspCysPro 63  
Db 769 CCGCGGCTCATCTCCACGAGACTGAGAGAGCTGCTCCCTGGAGGAGACCGAGAGTGTCCC 828  
Qy 64 TyrThrProLeuAspPheAsnProLeuAspLeuValTyrLeuSerThrThrLysLeuLeu 83  
Db 829 ACCGAGTGCACCTTCAGAAAGCGAATGTGGGCTCTGCTGATCATGATCCGAGTATGCTG 888  
Qy 84 SerAlaValAsnCysTyrPheLysCysMetValLeuAsnProTyrAsnGlnThrAspHis 103  
Db 889 CTGGACAAGTACTGCTGCTGACAGACCTGATGCTGACACAGACCGGCTTAGACCT 948  
Qy 104 ProGluCysLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGly 123  
Db 949 CCCCATTTCAAGCTCGAGCGCCGCGGAGATTGGCTCCCGGCACTTTTATGGCAGCG 1008  
Qy 124 TyrIleThrGlyProLeuSerThrValTyrLysGluTyrLeuLysTyrCysValGluPhe 143  
Db 1009 TAC-----TGGCTGTGGAACAAGCTGCTGGAACCTGGAGACCTGGCTGTATT 1050  
Qy 144 GlyIleGluAlaAsnAlaIleValAlaValProTyrAsnTyrArgLysLeuSerProThrLys 163  
Db 1051 GGATACGATACGATACATATGCTGCTGCGCGTACGATCGATGACATGCTTACCTGAT 1110  
Qy 164 LeuGluGluArgLysLeuTyrPheHisLysLeuLysLeuThrPheGluThrAlaLeuLys 183  
Db 1111 TTGGAGCGCCGAGACGATATCTTCCCAAGCTCAAGCTCAATCGAAGACATACGCGT 1170  
Qy 184 LeuArgGlyGlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArgTyr 203  
Db 1171 ATGACAGGTGAGAGACAGTCTGACGCGGCACTTCATGAGGCTCCAGGTCATCTTGTAC 1230  
Qy 204 PheLeuGluTyrPheArgLysLeuGluLeuAlaProLysHisTyrLeuLysTyrPheLysProL 223  
Db 1231 TTCATGAGTGGCTGAGCGCCGAGGATATGAGAGAGAGGTCCCACTGGGCTCAATGAC 1290

Qy 224 HisIleHisAlaTyrPheAlaValGlyAlaProLeuGluSerValGluAlaIleLys 243  
Db 1291 CATATTGAATCTTTGTGCAATTTCCGGCTCCATGCTGGAGTACCCCAAGACCTGGT 1350  
Qy 244 SerThrLeuSerGly-----ValThrPheGlyLeuProValSerGly 258  
Db 1351 GCTCTTCTTCTGCGAATAATGAAGATACCGTGCAGCTGAAGCGCATGGCTGTATGGA 1410  
Qy 259 ThrAlaArgLeuLeuSerAsnSerPheAlaSerLeu-----TyrLeuMetPro 275  
Db 1411 CTGGACGAGTCTTCTCTCGACGAGCGAGCTGCTGCGAACAATGAGGAGGAATT 1470  
Qy 276 PheSerLysAsnCysLysValGlyAsnAsnThrSerThrHisPheSerGlyIleAla 295  
Db 1471 GCTTCATGATTTCCCAAGGGGTGTAAAGCTATGAGTATGATCATTTGAGCCCT--- 1527  
Qy 296 LysLysAspLysArgValTyrHisCysAspGluGly----- 307  
Db 1528 -----GATGACGAGCCCGCGCAATGTCACCTTGGC 1560  
Qy 308 -----GluTyrGlnSerLys-TyrSerGlyTyr-- 316  
Db 1561 AACTCATCAAGTTCAAGAGTCTTGACCGAGTACTCTGTAGAACCTCACAAGAT 1620  
Qy 317 ----ProThrAsnIleIleAsnIleGluLeuProSerThrSerValThrGluThrAlaLe 335  
Db 1621 GAACCGGTGACTCTCGTATTTCTCAAGTCCGAGTGGTTGTGAACCGAGG 1680  
Qy 335 uValAsnMetThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArg 355  
Db 1681 G-----CTTACTCTTTGGAA-----TTGCCAAGACTCGAAAGAGGTGAGCGAAT 1728  
Qy 355 uLeuAlaAsp-----GlyThrLeuPheLysAlaIleGluAspTyrAspProAs 371  
Db 1729 GAGACGACGACTTCTACTCTGAGCAACCTCTGGAAG----- 1765  
Qy 371 PserLysArgMetLeuHisGlnLeuLysLeuTyrHisAspAspProValPheAsnPr 391  
Db 1766 -----CTGCTCTCCCC 1776  
Qy 391 oleuThrProTyrGluArgProProIleLysAsnValPheCysIleTyrGlyAlaHisLe 411  
Db 1777 AATG-----CCCCGATCTC-AAGATCTACTGTTTATGAGTCCGTA 1820  
Qy 411 uLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAsnTyr 431  
Db 1821 GATACCGAGCGAGCTACTACTACAGATGAGCCCAATCCCGAGCAGACCAATTGA 1880  
Qy 431 eIleThrAspIleIleTyrGluThrGluLysLeuValSerArgSerGlyThrValVa 451  
Db 1881 CGTCAAGTATGCTGGAACGACCTGATGTTGTATGAGGTACAG----- 1926  
Qy 451 IaAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSerLeuSer 471  
Db 1927 -----GCGCATGGAACCGTCTCTCTGTGACCATACAT 1961  
Qy 471 PCysLysAsnTyr 475  
Db 1962 GTGTACCGATGG 1974  
  
RESULT 8  
US-10-424-599-17331  
; Sequence 17331, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B

```

1 CURRENT APPLICATION NUMBER: US/10/424,599
2
3 CURRENT FILING DATE: 2003-04-28
4
5 NUMBER OF SEQ. ID NOS: 285684
6
7 SEQ ID NO 17331
8
9 LENGTH: 1433
10
11 TYPE: DNA
12
13 ORGANISM: Glycine max
14
15 FEATURE:
16
17 NAME/KEY: unsure
18
19 LOCATION: (1)..(1433)
20
21 OTHER INFORMATION: unsure at all n locations
22
23 FEATURE:
24
25 OTHER INFORMATION: Clone ID: PAT_MRI3847_115655C.1
26
27 US-10-424-599-17331

```

Alignment Scores:	
Pred. No.:	3, 27e-26
Score:	309.00
Percent Similarity:	72.55%
Best Local Similarity:	55.80%
Query Match:	10.42%
DB:	7
	Gaps: 2
	Length: 143
	Matches: 61
	Conservative: 13
	Mismatches: 18
	Indels: 10

US-09-651-651-5 (1-546) x US-10-424-599-17331 (1-1433)

QY ThrH<sub>1</sub>AphSeSGIyGLylalalalySlyAsPlyArGValrHsiScyASAPGIUGlu 307  
Db ACCCAAGCGCTCCGGGCAAGTGGGTCCACCA-----ACAATTCACtGTATGACGA 67  
QY GIUrYrgInSerLySrySrGIyTrpProthrAAsnIlelleAsnIleguileProSer 327  
Db GAGTTTAACACAACAACTTATCTGGGTGGCCACAAMATATCAACTTGAAATTCCTCA 127  
QY Thr-----SerValThrGIUthrAlaleuValAasMetThrSer 340  
Db ACTCGGCATTGTATGCATATCCTTATCTTCAGAAATACCTGAGGCCAACTGTCCAGC 187  
QY MetGIUCySGIyLeuBProThrLeuLeuSerPheThrAlaArgGIULeuAlaAspGIYThr 360  
Db ATGGAGTGGGAGTACTACTCATATATCTTTCTCAGCTCGGGAAATATCAGATGGCACCC 247  
QY LeuPheLYAlailEgLIUsApTYrAsPProAsPSerLyArGMetLeuHisiGlnLeuLys 380  
Db TTTTTTCAGGCAATTAAAGATTATGACCACAGACAGAAGGCTCTTGTATCCGTTAGAG 307  
QY LysLeu 382  
Db AGTTA 313

```

RESULT 9
US-10-321-802-25
; Sequence 25, Application US/10321802
; Publication No. US20030200563A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahn, Edgar B.
; APPLICANT: Camodu, Rebecca E.
; APPLICANT: Farnoud, Omolayo O.
; APPLICANT: Hall, Sarah E.
; TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
; FILE REFERENCE: Bbl486 US NA
; CURRENT APPLICATION NUMBER: US/10/321,802
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2398
; TYPE: DNA
; ORGANISM: Helianthus sp.
US-10-321-802-25

```

Alignment Scores:  
Pred. No.: 8.2e-23 Length: 2398

Score:	284.00	Matches:	142
Percent Similarity:	35.48%	Conservative:	84
Best Local Similarity:	22.29%	Mismatches:	258
Query Match:	9.58%	Indels:	153
DB:	6	Gaps:	28

US-09-651-651-5 (1-546) x US-10-321-802-25 (1-2398)

```

QY      14 VALLILEALVALPhePheLeuLeuCy-----GlyGlyArgThrLeuValaGluAap--- 30
Db      25 GTGGTTGTGTGTTTGTATTAAGTATGCCGCGCGTCTCCGCAAGTTTGTACGGA 354
QY      31 -----GluThrGluPheHis 35
Db      355 AGCGATATCCGAGCCGTTCCGAGATCTCCGAGATTAAAGTTTGAAGAAAGGTTTGA 414
QY      36 GLYAAPTYrSerLeuLeuSerGly-IleIleIleProGlyPheHisSerThrGlnLeuAr 55
Db      415 GCGGAA-----GCATCCCGGTGGTGTGTTGTGCGCGGAGATTGTACACCGGTGACTTGA 455
QY      55 GAlATPserIleLeuAapCyAapProTYrThrProLeuAapPheAapProLeuAapLeuVa 75
Db      466 GCTGTGGGAGAGCGACCAAGTGT-----ATGAGATGAGATTGCTTCCGAGAAAGCGCT 513
QY      75 ITrPLeuAapThr---ThrLeuLeuLeuSerAlaValaLancSTTPheLeuCyMetVa 94
Db      514 TTGGGGCGGTACGTTTGTGTGAGGATTATAAGAGCGCTTGCTGTGGTGTACAAATATGTC 573
QY      94 ILeuAapProTYrAsnGlnThrAapHisProGluCyValySerArgProAapSerGlyLe 114

```

Db	574	CGTAACAAACAAACCCGGATGATGCACCCGGGTATACGGGTCAGGCTGTCAATGACAT	633
Qy	114	USerLalLeIthrGluLeuAeApProGlyTYrIlePhcIyPProLeuSerThrValTPrLy	134
Db	634	TGTAACCTGTGACTACTCTTCGCTCCAGGGATATTT-----GTTGGGCG	675
Qy	134	SeGIuTrPLeuLySTrPCySValGIuPhGlyIleGIuLaIaAnaIaIeValaIaValaIpr	154
Db	676	TGTTTGTATGTCTAACTTGGACCGTGTGGGTATGAGAGAAATATGTATATGCTGTGC	735
Qy	154	oTYrAePTPrPaIGLeuSePProThrLyLeuGIuLrGaAePLeuTYrPhEhIaLyale	174
Db	736	ATATACCTGAGAACCTCGTTTCAAAACACAGAGGTAAAGACCAACACATGAGCGGAT	795
Qy	174	uLySeuThrPhEgIuThrAlaLeuLyLeuAryGIyGly---ProSerIleValIpheal	193
Db	796	AAAGAGCAATATAGAACTGATGGTTCCTACTAATGCGGAAAAAGCGGTATATATCC	855
Qy	193	aHisSerMeGIyAeAnaAnValPheArTYrPheLeuGIuTrPLeuAryLeuGIuIleal	213
Db	856	GGATTCAATGGGTGTATTACTACTTCTCGATTTCAATGAATGGTGGAG-----GC	906
Qy	213	aPProLyshIstYr-----LeuLySTrPLeuAePGLInHIsIleHIsIaIaTy	228
Db	907	ACCACACCAACATGGGTGGCGAGGTGACACAGATGGTGTGCTAAACACATAAAGCGGT	966
Qy	228	rPheAlaValGIyAlaPProLeuLeuGIy-----	237
Db	967	GATGAATATCGGTGAGCCATTTTTCAGGTGTCCCAAAAGCTGAGCGGCTTTCTCTGC	1022
Qy	238	-----SerValGIuAlaIeLySerThrPLeuSerGIyAla-----Th	250
Db	1027	GGAGCGTAAAGATATGTGCATCGTACGAGGCGCTTGCACACAGGTGTGTAGACTCGCATTT	1081
Qy	250	rPheGIyLeuPProValSerGIuGIyThrAlaArgLeuLeuSerAraenSerPheAlaSerSe	270
Db	1087	ATTTCAGATTCAACGTTAACACATGTATATGAGA--ATGACCGCCACATCGGATTCAAC	1144
Qy	270	rIeuTrPLeuMeTProPheSerLyAeAnCyLeuGIyAeAraenTrSerTrPThIaIePh	290
Db	1144	CATGTCTATGATACCG-----AAAGCGCGGACACACATTTGG-----	1188
Qy	290	eSerGIyGIyAlaIaIeLyAeAPLyAeArGIyValTyrHIsCyAeAPGIuGIuGIyTGTcl	310

```

Db      1181 -----GGGGGCTCAATTGGTCACCCGGAAGAGGTTATAGTCCAGAGAGAGTAACAATGG      ::::: 1236
Qy      310 nserLyseTserGlyTrpProthr-----                               318
Db      1237 AAAAAAGACACTGAATCTCCCAACCGTAAGATTTCTGCAAGTAAAGTAAACATGCAAA      ::::: 1296
Qy      319 -----AsnIleIleAsnIle-----GlnIleProserThrservalTh      331
Db      1297 TTATGAAAGGATGATATCTGTTCCGGAGAGATGACAGAGGCACCATCTTCAAGATCGA      ::::: 1356
Qy      331 rGluThrAlaLeu-----ValAsmMetThrsMetGluCys----- 343
Db      1357 GAGGATAGAAATTTAGAGGTGCTGTGAAGGGGATATCAATGTTCCAAACAATACATGTCCGGC      ::::: 1416
Qy      344 -----GlyLeuProThrIleuLeuSer----- 350
Db      1417 CGTGTGGACCGAATATACCATGATCATGGATTTGTGTGGAATCAAGGCTGTGGAGTACAA      1476
Qy      351 ----PheThrAlaArgGluLeuAlaAspGlyThrLeuPheValaIleGluAspTrpAs      369
Db      1477 GGTATATACAGCTGCGCGGAATCGTGAACTGCTGGAAGTTTCTCCAAAATATGATGGA      1536
Qy      369 pProAspSerLysArgMetLeuHisGlnLeuLysLeuTrpHisAspAspProValPhe      389
Db      1537 ACGGCGTAGTGCTCATTTTTCGTATGCTATGCTGCACAATTTG--GATGATCCAAAATA      1593
Qy      389 eAsnProLeuThrProTrpGluAspPro-----ProIleLys 401
Db      1594 TACACATTACAAAGATTTGGTCTTAACCCATTGGAGACAAAGTTACCAACGGCTCCAGACAT      1653
Qy      401 sAsnValPheCysIleTrpGlyAlaHisLeuLysThrGluValGlyTrpPhe----- 419
Db      1654 GGAGATCTATTCAAATGTATGAGATTTGGACTGCCAACCCAAAGAGCATTTGCTCAAACT      1713
Qy      420 -AlaProSerGlyLysProTrpProAspAsnTrpIleLeuThrAspIleIleTrpGluTh      439
Db      1714 CACACCTGCGGACAGAGTGTACATACCATCCAAAT-----GACAC 1755
Qy      439 rGluGlySerLeuValSerArgSerGlyThrValValAspGlyAsnAlaGlyProIleTh      459
Db      1756 GTCAGCAAAAGATAAAGCCGAGACGCGGTGTTAAANACCGG---GTTTATACGGTTGA      1812
Qy      459 rGlyAspGlyThrValProTrpHisSerLeuSerTrp-----CysLysAsnTrpLeuGly      477
Db      1813 CGGGGATGAAACAGTACAGACACTTAAGCGCGGTACATGTGCGCAAAAGGGTTGGCGTGG      1872
Qy      477 y-----ProLysValAsnIleThMetCysLysProGlnProGlnHisAsp-- 491
Db      1873 GAAAAACGGATTCAATCCTTCGGGAATTAACCTTATGTCAGGGAATACGATCACAAACC      1932
Qy      492 -----GlySerAspValHisValGluLeuAsnVal 501
Db      1933 TCCATCCAACTTTCTCGAGGGCCGGGGGACATCAAAAGCCGGGCCCATGTGATATTATGG      1992
Qy      501 lAspHisGluHisGlySerAspIleIleAlaAsnMetThrLysAlaProArgValLys-- 520
Db      1993 TAATTTCAGTTGATTGAAGATGTATATAAGATTGACAGCGGAGCCACGGGTGAAGAAGCT      2052
Qy      521 -----TyrIleThrPheTrpGluAspSerGluSerIle 531
Db      2053 TCGAGGTGATCAGGTGTACACTGTGTATATTGCAAGTGTCCGAGAAAGATC 2101

RESULT 10
US-10-321-802-23
; Sequence 23, Application US/10321802
; Publication NO. US20030200563A1
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Famedu, Omolayo O.
; APPLICANT: Hall, Sarah E.

```

```

TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
FILE REFERENCE: B81486 US NA
CURRENT APPLICATION NUMBER: US/10/321,802
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Microsoft Office 97
SEQ ID NO 23
LENGTH: 2700
TYPE: DNA
ORGANISM: Glycine max
US-10-321-802-23

Alignment Scores:
Pred. No.: 1,14e-22 Length: 2700
Score: 283.50 Matches: 119
Percent Similarity: 36.94% Conservative: 69
Best Local Similarity: 23.38% Mismatches: 219
Query Match: 9.56% Indels: 102
DB: Gaps: 20

US-09-651-651-5 (1-546) x US-10-321-802-23 (1-2700)
QY 44 11e11e11epprogiyPhealaserThrGlnLeuAgaLaItrpserilleuaspCysPro 63
DB 678 GTTTTGTGCCCGGAGATTGCATCGGGGCTTGAACTGGAGGGGTCACTGTGCT 737
QY 64 TyrThrProLeuaspPheasnProLeuaspLeuValItrpLeuaspThr--ThrLysLeu 82
DB 738 GAGGGGTGTTC-----AGMAAGCGTTGGGGGTGTACTTTTGGAGAGTCT 785
QY 83 LeuSerAlaValaenCysSTrPheLysCysMetValLeuaspProtyraenGlnThrAsp 102
DB 786 TATAAAGACCTTCATCTGGGTGGATCACATGTCATGGACATGAACAAGATTGGAT 845
QY 103 HisProGluCysLysSerArgProaspSerGlyLeuSerAlaIleThrGlnLeuaspPro 122
DB 846 CCACCAAGGCAATGAAGATTAGGCGCTGTCTGGACTTGTACTGCTGATTAATTGGCTGCA 905
QY 123 G1yTyrIleThrnglyProLeuSerThrValItrpLysGlnItrpLeuLysTrpCysValGlu 142
DB 906 GGATATCTTC-----GTTTGGGCACTCTTAATTTGCTAACTTGGGACGC 947
QY 143 PheGlyIleGlnAlaasnAlaIleValAlaValProtyraeTrpArgLeuSerProThr 162
DB 948 ATTGGTTATGAGAAAAAACTATGTACATGGCTGCATATGATTTGAGAAATPAGCATTTTCAG 1007
QY 163 LysLeuGlnGluAArgaspLeuTyrPheHisLysLeuLysLeuThrPheGlnThrAlaLeu 182
DB 1008 AACACTAGAGGTGAGGAGATCAAAACCTAAGTCGATTAATAAAGCAACATAGAACTTATGGTT 1067
QY 183 LysLeuAArgIyGly---ProSerIleValPheAlaHisSerMetGlyAsnAsnValPhe 201
DB 1068 GCTACTATAGTGTGAATAAGGCAAGTATATTCACATTCATCAAGGGGGCTTGTACTTC 1127
QY 202 ArgTyrPheLeuGlnItrpLeuAArgLeuGlnIleAlaProLysHisTyr----- 217
DB 1128 CTACATTTTATGAATGGGTGAA-----GCACACAGCTCCAAATGGGTGGGGGCA 1178
QY 218 ---LeuLysTrpLeuaspGlnHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeu 236
DB 1179 GGACCAGATTGTGTCTCCAAATATATTAAGGCAAGTGTAAACATTTGGTGCACATTTTTA 1238
QY 237 GlySerValGlnAlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSer 256
DB 1239 GGTGTTCACAGGCTATPAGCAAGGCTATTCTCAGCTAGGCGCAGAGATATGTGCTGTT-- 1295
QY 257 GlnGlyThrAlaArgLeuSerAsnSerPheAlaSer----- 270
DB 1296 -----GCCAGACAGATAGCTCCAGAGATTTTATAGATAAAGATCTGTTTGGATTCAA 1346
QY 271 -----LeuTrpLeuMetProPheSerLysAsnCys 280
DB 1347 ACCTTGCAACATGTATAGATAGATACCCGTACTTGGAGCTACAAAGTCAATATATACCA 1406

```

[illegible]

```

Db      907  -----AAGACCAATGATTCAGAGCTTCTTAAGAGTGTTCATGGGAACTCTGCGAAGTT 960
               :::::  |||  |||  |||  |||
Qy      321  -----IleAsnIle-----GluIle 325
               ::|||  |||
Db      961  AATCCAGAGCGCTGGAACTTTGAAAGATGSTATCTTTGGAAAAAGATGACGGAAGCT 1020
               ::|||
Qy      326  ProSerThrSerValThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeu 345
               :::::  :::::
Db      1021  CCGGCTTCCAATATATTGAGCAG-----ATGAATTCGCGTGAATGCTGTCCAAAGGTAAAT 1071
               :::::
Qy      346  ProThrIleLeuSerPheThrAla-----ArgGluIleAla 357
               |||  |||  |||
Db      1072  AATCTTGCCCATTTGGAAATACATCAATGCCGGAGTGTCTGGACAGAGTATCAGAAATTAGGG 1131
               |||  |||  |||
Qy      358  AspGlyThrIleuPheIleAlaIleGluAspIleArgPro----- 370
               |||  ::|||  |||
Db      1132  TGGGGGTGGAAAT-----AAGCAGATTTCAGACTCAAAAGCTTTCACCGCAGGCTTATCATTA 1188
               ::|||
Qy      371  -----AspSerLysArgMetLeuHisGlnLeuLysLysLeuThrHis-- 384
               :::::  :::::
Db      1189  GATCTTTTAACTTTGTGTCTCCAAAGGATGATGACGCTGTAGTGTTCATTTTCATTAATAT 1248
               ::|||  ::|||
Qy      385  -----AspAspProValPheAsnProIleuThrProTyrGluAspPro 398
               |||  |||  |||
Db      1249  GGAATTGCTGTAACCTTGATGATATCCAAATAATGGCCACTACAAAGTATTTGATCAAAACCC 1308
               |||  |||  |||
Qy      399  -----ProIleLysAsnValPheCysIleTyrGlyAlaHis 410
               ::|||  ::|||
Db      1309  TTGGAGACAAACATACACGATGCCGCTCAAAATGGAAATATTTTGCATGTATGAGAGTAGCC 1368
               ::|||
Qy      411  LeuLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAsnTyr 430
               ::|||  ::|||
Db      1369  ATTCCTACCGAAAGACATATGTCTATATAATTATTC-----CCACAGCAGCAGTGCATAT 1422
               ::|||
Qy      431  IleIleThrAspIleIleTyrGluThrGluLysSerLeuValSerArgSerGlyThrVal 450
               |||  |||  |||  |||
Db      1423  ATACCCTTTCAGATAGATGCCCTCAGCTAGAGGT-----GGGAGTAGAATATGCTGC 1473
               |||  |||
Qy      451  ValAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSerLeuSer 470
               ::|||
Db      1474  TTGAAGAGTGTGTTTACATATGTCGAATGTGACGAGACTGTTCCAGTCTTAAGTTCACAG 1533
               ::|||
Qy      471  Trp-----CysLysAsnTyrLeuGly-----ProLysValAsnIleThr 483
               ::|||  |||
Db      1534  TATATGTGTGCCAAAGCATGGGTGGAAAACTGCTTCAACCTTTGCGCAGCAAGACT 1593
               ::|||
Qy      484  MetAlaProGlnProGluHisAsp-----Gly 492
               ::|||
Db      1594  TACGTGAGAGAGTATAGTCATCTCCACCCTTCGAATCTCTCGAAGGACGAGGGCACACAG 1653
               ::|||
Qy      493  SerAspValHisValGluLeuAsnValAspHisGluHisGlySerAspIleIle 510
               |||  |||  |||  |||
Db      1654  AGTGTGTCGCCACGTGATATTAATGGGAACTTTGCTTAATAGGAGGATATTAATC 1707
               ::|||
RESULT 12
US-10-321-802-17
Sequence 17, Application US/10321802
Publication NO. US20030200563A1
GENERAL INFORMATION:
APPLICANT: Butler, Karlene H.
APPLICANT: Cahoon, Edgar B.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Famodu, Omolayo O.
APPLICANT: Hall, Sarah E.
TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
FILE REFERENCE: BBI486 US NA
CURRENT APPLICATION NUMBER: US/10/321,802
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 2479

```

US-10-321-802-17	US-09-651-651-5 (1-546)	US-10-321-802-17 (1-2479)
US-10-321-802-17	US-09-651-651-5 (1-546)	US-10-321-802-17 (1-2479)
Alignment Scores:	2,65e-22	Length: 2479
Pred. No.:	280.00	Matches: 129
Score:	35.60%	Conservative: 80
Percent Similarity:	31.98%	Mismatches: 252
Best Local Similarity:	9.44%	Indels: 126
Query Match:	6	Gaps: 23
DB:		
44	11e1lle1leprpog1yphela1aserthchlnleuarg1a1trpser1lleu1aaspCy63	
Db	351 GTGTTTGCCCGGAGATTGTGACCGGTGGGCTTGAAGTGGGAAGGCATCAAGT---	407
Qy	64 TyrthrPro1euaAphPheAenPro1euaAse1uVal1rpleuAsePthr---Th1r1e1eu 82	
Db	408 -----ATGATGGGCTTCTTATGGAAGAGGCTGGGGTGTACGTTTGGGAGAGTT 458	
Qy	83 LeuSer1aVala1aenCy1trpPhe1yCy1weVal1leuA1a1rPro1y1rAe1ng1nTh1rA1p 102	
Db	459 TATATAAGCCCTTCATGTTGGGTGACAAATATGTCCTGGACAAACAATCGGATGAT 518	
Qy	103 HisProgluCy1y1se1rA1rProA1sPse1rG1y1e1uSer1a1le1thrG1u1e1uA1pPro 122	
Db	519 CCGCGAGATTAACCGGTCAGACCCGTTAGTGAGACTTGTAGCTGTGACTACTTTGCCCG 578	
Qy	123 Gl1y1r1le1thrG1y1Pro1e1uSe1rTh1rV1l1rple1u1rple1u1s1rCy1e1Val1G1u 142	
Db	579 GGAATATTTT-----GTTTGGGCTGTTTGGATTGCTTAAGTGGCCGCT 620	
Qy	143 Phegl11legl1ua1a1e1u1a1le1vala1a1Val1rPro1y1rA1p1rA1rG1e1uSe1rPro1thr 162	
Db	621 GTTGATATGAAGAAAGAAATATGATATATGCTGCATATGACTGAGACCTCGTTTCAA 680	
Qy	163 Lys1e1uG1u1u1a1rA1p1e1u1rPhe1h1y1s1e1u1y1s1e1u1rPhe1G1u1Th1rA1leu 182	
Db	681 AACACGGAGGTAGACACCAACATTCGATGCTGATTAAGACCAATATAGACTAGTGGTT 740	
Qy	183 Lys1e1uA1rG1y1y1---Pro1e1r1le1Val1Phe1a1h1s1e1r1e1G1y1A1a1n1a1n1Val1Phe 201	
Db	741 GCTACAAATGGTGGGAAAAAGGCGGTTATTTCCCACTTCAATGAAGGCTGTTATCTACTTC 800	
Qy	202 Arg1y1rPhe1e1uG1u1rple1uA1rG1e1uG1u1le1a1rPro1y1h1s1y1r----- 217	
Db	801 CTGCATTTTCATGAATAATGGTGCAG-----GCACCAAGCTCCAAATGGGTGGAGGT 851	
Qy	218 ---Leu1y1r1rple1uA1sP1n1h1s1le1h1a1a1rPhe1a1a1G1y1a1a1rPro1e1u1e1u 236	
Db	852 GGACCAAGATGGGTGTGCTAAACATATCAAAACCGTAATGAACATTTGGTGACCATTTTTA 911	
Qy	237 Gl1y1e1rVal1G1u1a1le1y1s1e1rTh1r1e1uSe1rG1y1a1l1n1r----- 250	
Db	912 GGTGTCCCAAAAGCGTGAAGCCGGGCTTTCTTCGACAAAGCTAAAGATATTGCATCAAGTC 971	
Qy	251 -----Phe1y1e1uPro1v1a1Se1rG1uG1y 258	
Db	972 AGGCTCTTCGACCAACAGATATGTGACCTCGGATTTATTTTCAGATTTCAGAGCTTACAAAT 1031	
Qy	259 Thr1a1a1rG1e1u1e1uSe1rA1n1Se1rPhe1a1a1Se1rSe1r1e1u1rPle1uMetProPhe1Se1r1y1s 278	
Db	1032 ATAAATGAGA---ATGAGCCGACCAATGGATTCACATGCTTAATGATACCA----- 1079	
Qy	279 AenCy1y1y1e1G1y1A1sP1a1n1Th1r-----Se1rTP1n1h1a1Phe1Se1rG1y1y 293	
Db	1080 -----AAAGGGGGGACACCAATTTGGGGCGGTCTTAATTTGGTCTCCGAAAGCGGAT 1133	
Qy	294 Ala1a1a1y1y1e1A1p1y1A1rG1a1l1y1rH1s1Cy1a1e1rG1u1G1u1y1rG1n1Se1r1y1e1r 313	
Db	1134 ACCTCAAGTAAAGAAACATGGAATAAATGACACTGAATCTTCAACCAAAATAGAGCT 1193	

```

QY      314 SerGlyTyrProThrAsnIleIleAsn----- 322
DB      1194 GGAAGTGAAGAAATGTGAAGTAAACACGCAAAATTATGAGAGATGATCATTTGGGAGA 1253
QY      323 -----IleGluIleProSerThrSerValThrGluThrAlaLeu----- 335
DB      1254 GATGTAGCAGAGGACCATCTTTCAGAGATCGAGAGATGAATTGAGGGGTGCTGGAAG 1313
QY      336 ---ValAsnMetThrSerMetGluCys----- 343
DB      1314 GGTAAACAATGTTGCAAAACATACATCCGGGCGGTGGAACCGAATACCATGACATGGGA 1373
QY      344 -----GlyLeuProThrLeuLeuSer-----PheThrAlaArgGluLeuAlaAsp 358
DB      1374 TTTGTGTGATCAAGGCTGTTCAGAGATCAAGAGATATATACATGCTGGCAAAATTTGGAGT 1433
QY      359 GlyThrLeuPheLeuAlaIleGluAspTyrAspProAspSerLysArgMetLeuHisGln 378
DB      1434 ATGCTGGAAGTTTCTTCTCCAAAAATGATGGAACGGGCGAGTCTCATTTTCATATGTT 1493
QY      379 LeuLeuLeuLeuTyrHisAspAspProValPheAsnProLeuThrProTyrGluArgPro 398
DB      1494 ATAGCTGACAATTTG---GATGACCCAAATATCTCATTAACAATGATTGTTAAACCA 1550
QY      399 -----ProIleLeuAsnValPheCysIleTyrGlyAlaHis 410
DB      1551 TTAGACACAAGCTACCAAAACGCTCCAGACATGAGATCTATTCATATGATGGAGTTGGC 1610
QY      411 LeuLeuThrGluValGlyTyrTyrPhe-----AlaProSerGlyLysProTyrProAsp 428
DB      1611 ATCCCACTGAAGAAGACATATGTTTAAACTCACACCTGACAGACAGATGCTACATACCA 1670
QY      429 AsnTyrIleIleThrAspIleIleTyrGluThrGluGlySerLeuValSerArgSerGly 448
DB      1671 TTCCAAAATT-----GACACGCTCGCAAAAGATAAACGAGATGGG 1712
QY      449 ThrValValAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSer 468
DB      1713 TGTTTAAAGATGA---GTTTATACGGTGGACGAGATGAACAGATACCATTTAAAGC 1769
QY      469 LeuSerTyr-----CysLysAsnTyrPheGly-----ProLysValAsn 481
DB      1770 GCAGGCTACATGTGTCCAAAGGTTGGCGTGGGAAAAGTATGATTCATCTTGGGATTC 1829
QY      482 IleThrMetAlaProGlnProGluHisAsp----- 491
DB      1830 AAAACGTATGTTAGGAATACGATCAACATCTTCATCCAACTTCTTGGAGGCGCGGCG 1889
QY      492 ---GlySerAspValHisValGluLeuAsnValAspHisGluHisGlySerAspIleIle 510
DB      1890 ACCCAAAAGCGGGCTCACCGTGAATATTATGGTAAATTTCCAGTTAAATTGAAGATGTTATA 1949
QY      511 AlaAsnMetThrLysAlaProArgValLys-----TyrIleThrPhe 524
DB      1950 AAGCTTGACGCGGACGAGGTGAAGAACTGGGAGTGATCAGGTGATCACAGGTATA 2009
QY      525 TyrGluAspSerGluSerIle 531
DB      2010 TTCAGGTGTGTCGAGAAATTC 2030

```

# RESULT 13

```

US-10-437-963-38278/c
; Sequence 38278; Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

```

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38278
; LENGTH: 2395
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4192C.1
US-10-437-963-38278

```

Alignment Scores:

Pred. No.:	Length:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Gaps:
1,346-21	2395	274.00	34.15%	22.76%	7	26

```

QY      44 IleIleIleProGlyPheAlaSerThrGluLeuArgAlaTyrSerIleLeuAspCys--- 62
DB      2017 GTTTCGTCCCGGACATGCTACCGGCGCTCGAGCTCTGGAGGGGACACAGATCGCGC 1958
QY      63 -----ProTyrThrProLeuAsp 68
DB      1957 GAGGGGCTTCCCGAAGCGCTCTGGGCGGACGTTTCGCGAGCATGTATACAGACAT 1898
QY      69 PheAsnProLeuAspLeuValThrLeuAsp-----ThrThrLysLeuLeuSerAlaVal 86
DB      1897 TTTCTGTGCTTTGAT---GTTTGTGTGATGTTACAAAACATCAATGATATATGGGCT 1841
QY      87 AsnCysTyrPheLeuCysMetValLeuAspProTyrAsnGlnThrAspHisProGluCys 106
DB      1840 TTAATGCTGGTTGAACATATATGCTACATGACATGAACATGATGATTAACACAGGATTA 1781
QY      107 LysSerArgProAspSerGlyLeuSerAlaIleThrGluLeuAspProGlyTyrIleThr 126
DB      1780 AAGGTTCCGCGACGTACAGAGCTAGTGGCAGACATATTTGTTCTCGGTATTTT--- 1724
QY      127 GlyProLeuSerThrValTyrLysGluTyrLeuLysTyrCysValGluPheGlyIleGlu 146
DB      1723 -----GTTTGGGCTGTTTGAATTTAGCTCGATTTGATATGAA 1679
QY      147 AlaAsnAlaIleValAlaValAlaProTyrAspTyrArgLeuSerProThrLysLeuGluGlu 166
DB      1678 GAAAAGACCATATACATGCTGCTCATATGATTTGAAGTTATCTTTCCAGAACCTAGACT 1619
QY      167 ArgAspLeuTyrPheHisLeuLeuLysLeuThrPheGluThrAlaLeuLysLeuArgGly 186
DB      1618 CGTGATCAAACTTTGACGAGATATAAAGTAAACATTGAACCTCTGTGACCACTAATGGT 1559
QY      187 GlyProSer---IleValPheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeu 205
DB      1558 GGAATATAGGCTGTGTGATCCACATCTTAAGGGGTTCTCTATTTCTTGATTTTATG 1499
QY      206 GluTyrPheArgLeuGluIleAlaProLysHisTyr-----LeuLysTyr 220
DB      1498 AAGTGAGTTGAG-----GCTCTCTCTCCATACGAGGTGGTGGTGGTCCAAATTCG 1448
QY      221 LeuAspGlnHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuGlySerValGlu 240
DB      1447 TGTGCAAGCACATCAATCTGAATGAAATATTGGCGGACCTTTCTTAGAGTTCTTAAG 1388
QY      241 AlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGluGlyThrAla 260
DB      1387 GCTGTTCAGAGACTTTTCTCATCTGAAGCCAAAGATGTTGCTGTGCTAGACCATTTGCA 1328

```



Qy 261 ---ArgleuLeuSerAenSerphe-----Ala 268  
 Db 1327 CAGAGAGTCTAGACTGCTGCTTGAAGCTTACAGCCTTACGCCATTGATGCGTATG 1268  
 Qy 269 SerSerleuTrpLeuMetProPheSerLeuAenCysLeuAspAenThr-----285  
 Db 1267 ACCCGCAGCATGGGATTCACAGATGCTATGATGCTTAAAGGCGGTGACACCATTTGGGGA 1208  
 Qy 286 -----SerTrpThrHisPheSerGlyValAlaIaIys---LysAspLysArgValTyr 302  
 Db 1207 GATTGGATTGGTCTCCAGAAAGATGTTTGAAGTGAAGCTAAAGATGAATCGAAAAATC---1151  
 Qy 303 HisCysAspLeuLeuGlu-TyrGlnSerLysTyrSerGly-----TrpProThrAs 319  
 Db 1150 -----AATGATTCCTAGGTTTCTTAAGAGTCTTAACGGGAAGATGAGTTTCACAGAA 1097  
 Qy 319 nilelleAsnIleGluIleProSerThrSerValThrGlu-----332  
 Db 1096 CCTGTTAAGTATGAGAAATGTTCTCTTTCGTTAAGATGACAGAGGCTCCATCTTCA 1037  
 Qy 333 -----ThrAlaLeuValAenMetTh 339  
 Db 1036 GAAATTGACAGATAGATTTCTCTGTCGAGCTATTTCACATCTATTGACAGATGCT 977  
 Qy 339 rSerMetGluCysGlyLeuProThrLeuLeu-----SerPheThrAlaArgGluLe 356  
 Db 976 GTCAAGGCAATATATTGGCCACTCAAAATACATCATCCGGGATATATGACAGAGTAT 917  
 Qy 356 uala-----AspGlyThrLeuPheLeuValIleGluAspTyrAspProAspSerLysArg 374  
 Db 916 CATGAATTAGAGTGGCGCGA-ATTAAGCGAGTTGACAGACTTAACAGTTTACTGCTGG 858  
 Qy 374 gMeLeuHisGlnLeuLysLysLeu-----TyrHis 384  
 Db 857 CTCATTATAGATCTTCTGTTGTTGTTCTCCAGAGTATGACAGGCTGAAGTGTCA 798  
 Qy 384 g-----AspAspProValPheAsnProLeuThrProTyr 395  
 Db 797 CTTTCGATGCGATGCTGTCGACATCTGATGATCCAAAGTACCGCCACTCAAAATGACTG 738  
 Qy 395 pGluArgPro-----ProIleLysAenValPheCysIleTyr 407  
 Db 737 GTCAAAATCCCTTGGAAACAAATTAACCAATGACCTGAATGAATGAAATTTTCAATGTA 678  
 Qy 407 TGLValHisLeuLysThrGluValGlyTyrTyrPhe-----AlaProSerGlyLysPr 425  
 Db 677 TGGAGTTGGCATTCGACGAGAGAGCATATGCTATTAATTAAGCCCCACAGACAGACTG 618  
 Qy 425 oTyrProAspAenTrpIleIleThrAspIleIleTyrGluThrGluGlySerLeuValSe 445  
 Db 617 CTAT-----ATACCTTTTCAGATAGACGCTTCAGCTAGAGGT-----GG 579  
 Qy 445 rArgSerGlyThrValValAspGlyAsnAlaGlyProIleThrGluArgGluThrValPr 465  
 Db 578 CGATGAGAAATGCTCTGAAAGGCGGCTTACCTGTCTAATGATGATGAGACCGTACC 519  
 Qy 465 oTyrHisSerLeuSerTrp-----CysLysAsnTrpLeuGly-----Pr 478  
 Db 518 AGTTCTTATGTCAGAGATATATGTCGCGAAAGGCTGCGGAGGAAAAACAGCCTTCAACC 459  
 Qy 478 oLysValAsnIleThrMetAlaProGlnProGluHisAsp-----491  
 Db 458 TTTCGGCAGCAAGACCTGACGACAGATACAGCATTCCTCACCCCTCAAAATCTCTCGA 399  
 Qy 492 -----GlySerAspValHisValGluLeuAenValAspHisGlnHisGlySe 507  
 Db 398 AGCGAGGCGCAACCCAGAGTGTGCGCACGCTT-----368  
 Qy 507 rAspIleIleAlaAenMetThrLysAlaProArgValLysTyrIleThrPheTyrGluAs 527  
 Db 367 -GATATCATGGGGAATTTTGGCTTTAATGAGAGATTTATTCAGATTTGCTGCTGGGCGAAC 309  
 Qy 527 pSerGluSerIleProGlyLysArgThrAlaValTrpGluLeu 541

Db 308 TGGTGAAGAGCTTGGCGGTGACCAAGTATTCATTCAGGAATTTG 266  
 RESULT 14  
 US-10-425-115-55049  
 ; Sequence 55049, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ. ID NOS: 369326  
 ; SEQ. ID NO 55049  
 ; LENGTH: 2700  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRF4577\_150201C.1  
 US-10-425-115-55049  
 Alignment Scores:  
 Pred. No.: 2,48e-21 Length: 2700  
 Score: 272.50 Matches: 132  
 Percent Similarity: 33.11% Conservative: 69  
 Best Local Similarity: 21.75% Mismatches: 262  
 Query Match: 9.19% Indels: 144  
 DB: 8 Gaps: 22  
 US-09-651-651-5 (1-546) x US-10-425-115-55049 (1-2700)  
 Qy 44 ILeIleIleProGlyPheAlaSerThrGlnLeuAlaIleTrpSerIleLeuAspCys---62  
 Db 561 GTCTTCGTCCCGGACATGTCACCGGGGCTGAGTTATGAGAGGCGACCACTGGGCC 620  
 Qy 63 -----ProTyrThrProLeuAspPheAsnPro 71  
 Db 621 GAGGGGCTCTCCGCAAGCGGCTATGAGGGCGGACATTTGCTGATGACGATACAGAGACCT 680  
 Qy 72 LeuAspLeuValTrpLeuAspThrThrLysLeuLeuSerAlaValAenCysTrpPheLys 91  
 Db 681 CTA-----TCTCGGCTGGA 695  
 Qy 92 CysMetValLeuAspProTyrAsnGlnThrAspHisProGluCysLysSerArgProAsp 111  
 Db 696 CATATGTCATTGGACAAATGAAGTGAATTAGACAAACCTGAAATTAAGGTGACGGCGGTC 755  
 Qy 112 SerGlyLeuSerAlaIleThrGlnLeuAspProGlyTyrIleThrGlyProLeuSerThr 131  
 Db 756 ACAGGCGCTGGTGGACAGACATTTTGTCTCCGATATTTT-----797  
 Qy 132 ValTrpLysGluTrpLeuLysTyrCysValGlnPheGlyIleGluAlaAenAlaIleVal 151  
 Db 798 GTTTGGGCTGCTTAATTTGCAATTTTGAACGATTTGATGAGATGAAGAAAGACCATGTAC 857  
 Qy 152 AlaValProTyrAspTrpArgLeuSerProThrLysLeuGluGluArgPheLysTyrPhe 171  
 Db 858 ATGCGTCATATGATTTGAGAGTTATCTTTCCAGAAACCTAGAGTCCGTGATCAAACTTTG 917  
 Qy 172 HisLysLeuLysLeuThrPheGlnThrAlaLeuLysLeuArgGlyIleProSer---1le 190  
 Db 918 AGCGAATTAAGACCAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 977  
 Qy 191 ValPheAlaHisSerMetGlyAsnAenValPheArgTyrPheLeuGluTrpLeuArgLys 210  
 Db 978 GTATATCCACACCTCATAGGGGCTCTATTTTTCATTTTATGAATGAATGGTGAATG 1034  
 Qy 211 GluIleAlaProLysHisTyr-----LeuLysTrpLeuAspGlnHisIle 225

```

Db      1035 -----GCACCTTCCTCCCATGGGGGGTGGCGGTGTCACAGACTGGTGTGACGAAGCATATT 1088
               |||
               |||
               |||
Oy      226 HisAlaTyPheAlaValGlyAlaProLeuLeuGlySerValGluAlaIleLeuSerThr 245
               |||
               |||
               |||
Db      1089 AAGAGTGAATGAATATGTGAGAGACCTTCTTAGGATGCTTAAGAGCTGTCTGGCCTT 1148
               |||
               |||
               |||
Oy      246 LeuSerGlyValThrPheGlyLeuProValSerGlyThrAla---ArgLeuLeuSer 264
               |||
               |||
               |||
Db      1149 TTCTCATCTGAAAGCCAAAGATGTGCGCTTGACGATAGCTATGCTGATGCTTGGAC 1208
               |||
               |||
               |||
Oy      265 AasnSerPhe-----AlaSerSerLeuTrpLeu 273
               |||
               |||
               |||
Db      1209 TCTGATTTCTTGGGCTTCAAACTTTGCCCATTTAGTCGATAGACCCGAACATGGAT 1268
               |||
               |||
               |||
Oy      274 MetProPheSerLeuValSerCysGlyAlaAspAsnThrSerTrpThrIlePheSerGly 293
               |||
               |||
               |||
Db      1269 TCACAAATTTCAATGATGATCTTAAGGTGGTATACAAATTTGG-----GGAAAT 1316
               |||
               |||
               |||
Oy      294 AlaAlaValLeuAspLeuValArgValTyriLeuCyAspGluGluGlyTrpGlnSerTy 313
               |||
               |||
               |||
Db      1317 CTGGATTTGGTCTCCCAAGATGGCATTTGAATGTAAGACTAAGACAAAGCCAAATGAT 1376
               |||
               |||
               |||
Oy      314 SerGly-----TrpProThrIleAsnIleLeuIleGluIleProSerThrSerValThr 331
               |||
               |||
               |||
Db      1377 ACTGGGCTTCTAAGATAGCAATGGGGAATATGAGAGTTCAACCTGATTAAC 1436
               |||
               |||
               |||
Oy      332 GluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeuLeuSer--- 350
               |||
               |||
               |||
Db      1437 TATGAAAGCTGGTATCTTTGGTAAAGATGTAGCAAGACCTTTCAGAGAGTGTAG 1496
               |||
               |||
               |||
Oy      351 ----PheThrAlaArgGluLeuValAspGlyThrLeu----- 361
               |||
               |||
               |||
Db      1497 CAGATTAAGATTTTGTGATGCTGTTAAAGTAATATATGATGCCCATTTCAATACGTATGC 1556
               |||
               |||
               |||
Oy      362 -----PheLeuValIleGluAsp 367
               |||
               |||
               |||
Db      1557 CCGGAGATCTGGACAGATATACGAATTAAGATGGGTGAATTAAGACAGTCGACAGAC 1616
               |||
               |||
               |||
Oy      368 TyrAspPro-----AspSerValArgMet 375
               |||
               |||
               |||
Db      1617 TACAAAGTTTACACTGCTAGTCTGTATTAAGACCTTTCATTTGTGTTGCTTCCAAAGATG 1676
               |||
               |||
               |||
Oy      376 LeuHisGluLeuValLeuValLeuValHis-----AspAspProVal 388
               |||
               |||
               |||
Db      1677 ATGCAGCGTGAATAATGTCATTTGCATATGAAATTCGATTAAGTGAATGATGCCGAAA 1736
               |||
               |||
               |||
Oy      389 PheAsnProLeuThrProTrpGluArgPro-----ProIle 400
               |||
               |||
               |||
Db      1737 TACCAACATTACAAATATGGTCAAAACCTTGGAAACGAAGTTACGAATGCTCTGAC 1796
               |||
               |||
               |||
Oy      401 LysAsnValPheCysIleTyriGlyAlaHisLeuLeuThrGluValGlyTyriTrpPhe--- 419
               |||
               |||
               |||
Db      1797 ATGGAATATATTTCCATGACGATAGAGATGCTTCTTAAGAAAGGCAATATGCTTAAG 1856
               |||
               |||
               |||
Oy      420 ---AlaProSerGlyLysProTyriProAspAsnTrpIleLeuThrAspIleIleTyriGlu 438
               |||
               |||
               |||
Db      1857 TTGGCCCCACAGGACAAATGTAT-----ATACCTTCCGAATTGACACTCG 1904
               |||
               |||
               |||
Oy      439 ThrGluGlySerLeuValSerArgSerGlyThrValValAspGlyAsnAlaGlyProIle 458
               |||
               |||
               |||
Db      1905 GCGTAAGGC-----GGGAGAGAAATATAGCTGTTAAAGGGGTGTTCCTTAAGCC 1955
               |||
               |||
               |||
Oy      459 ThrGlyAspGluThrValProTyriHisSerLeuSerTrp-----CysValAsnTrpLeu 476
               |||
               |||
               |||
Db      1956 GATGTGATGAATACTGTCTTCAAGTCTTAAAGTCGCGGCTACATGTGTCAAAAGATGGCGGT 2015
               |||
               |||
               |||
Oy      477 Gly-----ProLysValAsnIleThrMetAlaProGlnProGluHisAsp 491
               |||
               |||
               |||
Db      2016 GGCAAACTCGTTTCAACCTGCGGACGACAGACTTACGTGACAGATATACGCCATTCA 2075
               |||
               |||
               |||
Oy      492 -----GlySerAspValHisValGluLeuAsn 500
               |||
               |||
               |||

```

```

Db      2076 CCACCTCTACTCTCTGGAAGGACAGGACCCAGACGGTGCATGTGATATATG 2135
Oy      501 ValAspHisGluHisGlySerAspIleIle-----AlaAsn-MetThr 515
               |||
               |||
               |||
Db      2136 GGGAACTTTGCTCTTATTAAGACGCTCATACGAATAGCTGTGGGCAACGGTGAAGA 2195
               |||
               |||
               |||
Oy      515 AlaProArgValLeuTyriIleThrPheTyriGluAspSerGluSerIleProGlyLysArg 535
               |||
               |||
               |||
Db      2196 ATTTGGCGCATCAGGTTTATTCAGATATATTCAAAGTGTGACAGAAAAATCAATTGAAA 2255
               |||
               |||
               |||
Oy      535 GThrAlaValTrpGluLeu 541
               |||
               |||
               |||
Db      2256 TTGTAACTATGGAAGTT 2274
               |||
               |||
               |||
RESULT 15
US-10-321-802-35
/ Sequence 35, Application US/10321802
/ Publication No. US20030200563A1
/ GENERAL INFORMATION:
/ APPLICANT: Butler, Karlene H.
/ APPLICANT: Cahoon, Edgar B.
/ APPLICANT: Cahoon, Rebecca E.
/ APPLICANT: Farnodu, Omolayo O.
/ APPLICANT: Hall, Sarah E.
/ TITLE OF INVENTION: Phospholipid:diacylglycerol Acetyltransferases
/ FILE REFERENCE: B01486 US NA
/ CURRENT APPLICATION NUMBER: US/10/321,802
/ CURRENT FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 35
/ LENGTH: 4093
/ TYPE: DNA
/ ORGANISM: Glycine max
US-10-321-802-35

Alignment Scores:
Pred. No.: 1,29e-20 Length: 4093
Score: 269.00 Matches: 116
Percent Similarity: 36.20% Conservative: 69
Beet Local Similarity: 22.70% Mismatches: 219
Query Match: 9.07% Indels: 108
DB: 6 Gaps: 19

US-09-651-651-5 (1-546) x US-10-321-802-35 (1-4093)
Oy      44 IleIleIleProGlyPheAlaSerThrGluLeuArgAlaTrpSerIleLeuAspCysPro 63
               |||
               |||
               |||
Db      615 GTTTTGTGCCCCGGGATTTGTCACTGGGGGCTTGAACGTGGAGGTCACCTGTGTCT 674
               |||
               |||
               |||
Oy      64 TyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThr---ThrLysLeu 82
               |||
               |||
               |||
Db      675 GAGGGGTGTTCT-----AGAAACCTTATGGGGTGTACCTTGGAGAAAGTT 722
               |||
               |||
               |||
Oy      83 LeuSerAlaValAsnCysIleTrpPheLysCysMetValLeuAspProTyriAsnGlnThrAsp 102
               |||
               |||
               |||
Db      723 TATTAAGAAGCTTCATGCTGGGTGATCATCATGTCACTGCAACATGAAGAAACAGATTGAT 782
               |||
               |||
               |||
Oy      103 HisProGluCysLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeuAspPro 122
               |||
               |||
               |||
Db      783 CCACCAAGGATGAAGATTAGGCTGTCTGTGACCTTGACCTTGATTAAGTTACTTGTCTGCA 842
               |||
               |||
               |||
Oy      123 GlyTyriIleThrGlyProLeuSerThrValTrpLysGluTrpLeuLysTrpCysValGlu 142
               |||
               |||
               |||
Db      843 GGATACCTTT-----GTATGGCAGTGTATATGCTTAATCTTGACGCC 884
               |||
               |||
               |||
Oy      143 PheGlyIleGluAlaAsnAlaIleValAlaValProTyriAspTrpArgLeuSerProThr 162
               |||
               |||
               |||
Db      885 ATTTGGTATGAAGAAAAAATATATGATCATGCTGATATGATTTGAGATATACATTTCAG 944
               |||
               |||
               |||
Oy      163 LysLeuGluGluArgAspLeuTyriPheHisLysLeuValLeuThrPheGluThrAlaLeu 182
               |||
               |||
               |||
Db      945 AACACTGAGGTGAGGATCAACACTAAGTCCGATTAAGAAAGCAATGAACTTATGTGT 1004
               |||
               |||
               |||

```

QY 183 LysLeuArgGlyGly---ProSerIleValPheAlaHisSerMetGlyAsnAsnValPhe 201  
Db 1005 GCTACTTAATGTGGTAATAAGGAGGTATATATCCACATTCATGAGGGGCTTGACTTT 1064  
QY 202 ArgTyrPheLeuGluThrLeuArgLeuGluIleAlaProLysHisTyr----- 217  
Db 1065 CTTCATTTTATGAAGTGGTTGAA-----GCACCACTCCCACTGGTGGTGAGGA 1115  
QY 218 ---LeuYerThrLeuAspGlnHisIleAlaTyrPheAlaValGlyAlaProLeuLeu 236  
Db 1116 GGAACCAAGATTGGTCTCCACATATATAAGCACTTGAACATGGTGAGCACTTTTAA 1175  
QY 237 GlySerValGluAlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSer 256  
Db 1176 GGTGTTCCCAAGGCTATAGCAGGGCTTTCTCAGCTGAGGCCGGGATATTGCTGTT--- 1232  
QY 257 GluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSer----- 270  
Db 1233 -----GCTAGCAATAGCTCCAGAGTTTATAGATAACGATCTGTTCCGATTCAA 1283  
QY 271 -----LeuThrLeuMetProPheSerLysAsnGly 280  
Db 1284 ACATTGCAACATGTAATGAAGATGACCCGTACTTGGAAGCTCAACATGTCATGATCCA 1343  
QY 281 LysGlyAspAsnThrSerTrpThrHisPheSerGlyGlyAlaAlaLysAspLysArg 300  
Db 1344 AGAGGAGAGATATATATG-----GCTGCTCTGATGTTGTCACCGAGAGAA 1391  
QY 301 ValTyrHisCysAspArgGluGluGlyTyrGlnSerLys-----TyrSer 314  
Db 1392 GGCATATCACCTAGCCAGAGAAAGCACAGCAATACATACTCAGTTGAAAGACACGAA 1451  
QY 315 GlyTrpProThrAsnIleLeuAsnIle----- 323  
Db 1452 ACAATCAACACAAATTTGTTCACTATGGAAGAAATGATATCTTGGCAGAGATGTGGCC 1511  
QY 324 -----GluIleProSerThr-----SerValThrGluThrAla 334  
Db 1512 GAGGCACACTCCCTGAGATTGATGATGATGCTCCGGGGTCTATCAAGGTCGCGAGT 1571  
QY 335 LeuValAsnMetThrSerMetGlu-CysGlyLeuProThr----- 347  
Db 1572 ATTGCAATATACACTGTGCGATGTGGAGTGAATGCAATGCAATGCGATTGGAAGA 1631  
QY 348 -----LeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLysAl 364  
Db 1632 GTGAGACAGCTTGTGAACATTAAGTTTACACAGCTGGCT-----CAGTC 1676  
QY 364 aileGluAspTyrAspProAspSerLysArgMetLeuHisGlnLeuLysLysLeuTyrHis 384  
Db 1677 GTTGAACCTCTTCAATTT-GTTGCTCCAAAGATGATGCTGCTAGTGTCTCATTTCTC 1735  
QY 384 s-----AspAspProValPheAsnProLeuThrProTrpGluAr 397  
Db 1736 TTATGAAATTGCTGACATTTGATGACCCATAAATATATATCATCAACAAGTATGTCAAA 1795  
QY 397 gPro-----ProIleLysAsnValPheCysIleTyrGlyAla 409  
Db 1796 CCCCCTTGGAACAAATTAACCAAAATGCTCTGATATGGAATCTTCTCTATATAGAGCT 1855  
QY 409 aHisLeuLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAspAs 429  
Db 1856 TGGCTTACCTACTGAAGATCTTATATTTACAGTTAACT-----CCCTTGGCCGAGTG 1909  
QY 429 nTrpIleThrAspIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyThr 449  
Db 1910 TTACATTCCTTTTGAAATTGACACCAACAAGATGCTGT-----ACCGATGAAGATAG 1963  
QY 449 rValValAspGlyAsnAlaGlyProIleThrGlyAspGlyThrValProTyrHisSerLe 469  
Db 1964 CTGCTGCAAGGTGAGTCTACACTGTGATGGGATGAGACTGTGCGGTTCTTAAGTTTC 2023

QY 469 userTrp-----CysLysAsnTrpLeuGly 477  
Db 2024 AGGCTTCATGTGTGCTTAAAGTTGGCGCGGA 2054  
Search completed: November 22, 2005, 14:08:37  
Job time : 866 secs

THIS PAGE BLANK (USPFO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 22, 2005, 11:10:45 ; Search time 131 Seconds  
(without alignments)  
547,842 Million cell updates/sec

Title: US-09-651-651-5

Perfect score: 2965  
Sequence: 1 MGNASKSVASFTVIAVFL.....DSEISPGKRTAVELDXSGY 546

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3196817 seqs, 65720914 residues

Total number of hits satisfying chosen parameters: 6393634

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-Q=/cgn2\_1/USFTO.spool/US09651651/runat.18112005.101832\_20333/app.query.fasta\_1.711  
-DB=Published.Applications\_NA\_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdt -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09651651 @CGN 1.1 117 @runat.18112005.101832\_20333  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_New:\*

1: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	104.5	3.5	4333	1	US-10-131-826A-233
2	85.5	2.9	2806	1	US-10-131-826A-201
3	85	2.9	2906	1	US-10-131-826A-367
4	83.5	2.8	26000	1	US-10-949-720-391
5	82	2.8	7512	8	US-10-647-956A-7
6	80.5	2.7	1455	8	US-11-147-725-3
7	80.5	2.7	1991	8	US-11-147-725-1
8	79	2.7	2310	1	US-10-957-569-11

9	78.5	2.6	1023	1	US-10-467-962B-78	Sequence 78, Appl
10	78	2.6	2739	1	US-10-519-599-3	Sequence 3, Appl
11	77.5	2.6	9719	9	US-11-042-988-10	Sequence 10, Appl
12	77.5	2.6	48763	1	US-10-663-794-3	Sequence 3, Appl
13	77	2.6	1287	7	US-11-074-176-247	Sequence 247, Appl
14	76.5	2.6	1689	7	US-11-021-441-8	Sequence 8, Appl
15	76	2.6	1824	9	US-11-085-864-1	Sequence 1, Appl
16	75.5	2.5	1856	9	US-11-082-389-435	Sequence 435, App
17	75	2.5	1299	1	US-10-689-742-143	Sequence 143, App
18	74.5	2.5	2745	9	US-11-077-550-119	Sequence 119, App
19	74.5	2.5	2835	9	US-11-077-550-119	Sequence 121, App
20	74.5	2.5	2853	9	US-11-077-550-117	Sequence 117, App
21	74	2.5	2449	1	US-10-510-386-47	Sequence 47, Appl
22	73.5	2.5	2498	1	US-10-131-826A-483	Sequence 483, App
23	73.5	2.5	34000	7	US-11-102-978-3	Sequence 3, Appl
24	72.5	2.4	1881	1	US-10-512-184-24	Sequence 24, Appl
25	72.5	2.4	1989	7	US-11-021-441-10	Sequence 10, Appl
26	72.5	2.4	2030	1	US-10-652-893-3	Sequence 3, Appl
27	72.5	2.4	3003	9	US-11-038-284-12	Sequence 12, Appl
28	72.5	2.4	3003	9	US-11-038-284-39	Sequence 39, Appl
29	72.5	2.4	3327	7	US-11-021-441-44	Sequence 44, Appl
30	72	2.4	1869	7	US-11-021-441-34	Sequence 34, Appl
31	72	2.4	1875	1	US-10-131-826A-359	Sequence 359, App
32	72	2.4	2361	1	US-10-467-962B-102	Sequence 102, App
33	72	2.4	3552	7	US-11-021-441-43	Sequence 43, Appl
34	71.5	2.4	1968	1	US-10-131-826A-163	Sequence 163, App
35	71.5	2.4	2623	1	US-10-131-826A-63	Sequence 63, Appl
36	71	2.4	2737	7	US-11-021-441-12	Sequence 12, Appl
37	71	2.4	4372	1	US-10-131-826A-55	Sequence 55, Appl
38	70.5	2.4	4372	1	US-10-131-826A-53	Sequence 53, Appl
39	70	2.4	43948	1	US-10-949-720-393	Sequence 393, App
40	69.5	2.3	1462	9	US-11-012-762-11	Sequence 11, App
41	69.5	2.3	2477	1	US-10-131-826A-331	Sequence 331, App
42	69.5	2.3	2531	9	US-11-038-284-37	Sequence 37, Appl
43	69.5	2.3	2916	1	US-10-131-826A-69	Sequence 69, Appl
44	69.5	2.3	3033	9	US-11-038-284-14	Sequence 14, Appl
45	69.5	2.3	3033	9	US-11-038-284-41	Sequence 41, Appl

#### ALIGNMENTS

RESULT 1  
US-10-131-826A-233  
; Sequence 233, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C128  
; CURRENT APPLICATION NUMBER: US/10/131.826A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113

```
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059115
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 233
/ LENGTH: 4333
/ TYPE: DNA
/ ORGANISM: Homo Sapien
US-10-131-826A-233

Alignment Scores:
Pred. No.: 15          length: 4333
Score: 104.50         Matches: 60
Percent Similarity: 34.17%      Conservative: 35
Best Local Similarity: 21.58%    Mismatches: 80
Query Match: 3.52%             Indels: 103
DB: 1                      Gaps: 17

US-09-651-651-5 (1-546) x US-10-131-826A-233 (1-4333)
Qy 271 LeuTrpLeuMetProPheSerIlyAsnCylysglyAspAsnThrSerTrpThrHisPhe 290
    |||||  ::::|  |||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 155 CTCGGGCTCTGCTGTTTCATGCAGTCTTG-----TGGCTCAAGT 196
Qy 291 SerGlyGlyAlaAlaIlyValIyValGlyValHisCyAspGluGluGlyTrpGln 310
    ::::|  |||||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 197 ACTGATGGAGCCCT-----CGAGTCTACTGCGGATCGCGGATGTGC 244
Qy 311 SerIyTrSerGlyTrpProThrAsnIleIleAsnIleGluIleProSerThrSerVal 330
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 245 TGGAACTATGCTCCCAAGGAGAAATGTATACGACACGCTTGGACAGTGCATTA 304
Qy 331 ThrGluThrAlaLeuValAsnMetThrSerMetGlyCyAspIlyLeuProThrLeuSer 350
    ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 305 GTGGCTTCCAGCTTTTAAAGTGTGACAGAACCGGATGG----- 346
Qy 351 PheThrAlaArgGluLeuAlaAspGlyThrLeuPheIyAlaIle---GluAspIlyAsp 369
    ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 347 -----GGAACTTACAGAGCCATCTATTAAGATACAG 382
Qy 370 ProAspSerIyAspMetLeuHisGlnLeuIlyLeuIyTrpHisAspAspProValPhe 389
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 383 GATGACTCA-----TACACAGATGAAGTGGCC 409
Qy 390 AsnPro-----LeuThrProTrpGluArgProProIleIyAsnValPhe 404
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 410 CACCCGTGCTGTTGGGCTTCTGGGGCCAGTGTGAGGCTGAAGTGGGAGTGCAT 469
Qy 405 CysIleIyGlyAlaHisIleLeuIyThrGluValGlyTrpTrpPheAlaProSerGlyIy 424
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 470 CTATT-----CACCTGAAGAAT-----TTTGCCACT-----CGT 499
Qy 425 ProIyTr-----ProAspAsnTrpIleIleThrAspIleIleIyGly-----Thr 439
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 500 CCTTATACCATCCACCTCATGT-----CTCTTACGAGAGACTCT 544
Qy 440 GluGlySerLeuValSerArgSerGlyThrValValAspGlyAsnAlaGlyProIleThr 459
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 545 GAAGTTCCTCATACCA-----GATGGCTCTCTGGGCCACATGAAA 586
```

```
Qy 460 GlyAspGluThrValPro-----IyHisSerLeuSerTrp----- 471
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 587 GCTGATGACTGTGTCCCGGGGGGAGCCATATCTAACACTGACCATTCAGAGAGCC 646
Qy 472 -----CylyAsnTrp----- 475
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 647 CATGCACCCACCGATGCTGACCCAGCGCTGCACCTGATCTACCATTCATGTAGAT 706
Qy 476 -----LeuGlyProIyValAsnIleThrMechAlaPro 486
    ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 707 GCTCCACGACGATTCGAACTGCGCTTAATTTGGGCTCTCATCACC-----TGTA 757
Qy 487 GlnProGluHisAspGlySerAspValHisValGluLeuAsnValAspHisGlu 504
    ::::|  |||||  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|  ::::|
Db 758 AGAGAGCCCTGATGGAACTCCCTCTCAAGCCAGATGTAGACCATGAT 811

RESULT 2
US-10-131-826A-201
/ Sequence 201, Application US/10131826A
/ Publication No. US20050245730A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C128
/ CURRENT APPLICATION NUMBER: US/10/131,826A
/ CURRENT FILING DATE: 2002-04-24
/ PRIOR APPLICATION NUMBER: 60/049911
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: 60/056974
/ PRIOR FILING DATE: 1997-08-26
/ PRIOR APPLICATION NUMBER: 60/059113
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059115
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 201
/ LENGTH: 2806
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 2157
/ OTHER INFORMATION: unknown base
```

US-10-131-826A-201

## Alignment Scores:

Pred. No.:	606	Length:	2806
Score:	85.50	Matches:	100
Percent Similarity:	32.67%	Conservative:	48
Best Local Similarity:	22.08%	Mismatches:	157
Query Match:	2.88%	Indels:	149
DB:	1	Gaps:	21

US-09-651-651-5 (1-546) x US-10-131-826A-201 (1-2806)

```
QY      83 LeuSerAlaValAlaValLeuSerTrpPheLeuValLeuSerProTyrAsnGlnThrAsp 102
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      752 CTGCTGATTAACAATGCTGCG-----AGCATGCTTAAGCCTTAATAATACAACTGCGG 802

QY      103 HisProGluCysIysSerArgProAsp-SerGlyLeuSerAlaIleThrGluLeuAspTrp 122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      803 -----ACATTAAAGCCTGACGACAGAGGCTGTGGTTG-----CTGCCACC 841

QY      122 OGlyTyrIleThrGlyProLeuSerThrValITrplyGluTrpLeuLys----- 138
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      842 CGGCTGATAGTGTCTCTTTTCTTGGAATGTGCCCCAGGGCTGAAGCGCATGTGCT 901

QY      139 -----Trp-Cys-ValGluPheGlyIleGluAlaAsnAlaIleValAla 152
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      902 TCCTTTGTCAACGACGTGGCTGCTGCTGAAGCTTTGCAAAAGCGACCTGATGTGACCC 961

QY      153 ValPro----- 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      962 TGCCCGCGCAATGCATGTTGTCTTCTTTCAAAGGGAACTTTTGCATTCATTTGCCACC 1021

QY      155 -----TyrAspTrpArgLeuSer-----ProThrLysLeuGluGluArgSP 168
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1022 TCAGAGATGCTTACGATATGAGAAAGGCAAGTTTCCCGTCGATTAAGAAATGTGTGAC 1081

QY      169 LeuTyrPheHisIysLeuLysLeuThrPheGluThrAlaLeuLysLeuArgGlyPro 188
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1082 TCATTTGTGAGCTGGAGCAGGTGGCTTAAAGAACTTCATTAGAGCTTT--GGATGCACA 1139

QY      189 SerIleValPheAlaHisSerMetGlyAsnAsnValPheArgGlyTrpPheLeuGluTrpLeu 208
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1140 CAGATCTGTTTCTCAGAAAAATAGTCTGTACGAAACCGAGTGGAGCATCTCTGCG-- 1196

QY      209 ArgLeuGluIleAlaProLysHisTyrLeuLysTyrPheLeuSerGlnHisIleHisAlaTyr 228
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1197 -----CCACATTTGAGAAAG----- 1211

QY      229 PheAlaValGlyAlaProLeuLeuGlySer-ValGluAlaIleLysSerThrLeuSerG1 248
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1212 ---GTGGTGGCTGCTCCCTGCTGCTCATCTCAGAGAGGCAAAATCAGTCCAGCCTCTCC 1268

QY      248 ValIThrPheGlyLeuProValSerGluGlyThrAlaArgLeuLeuSerAsnSerPheAl 266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1269 CACCATCTTCCCTGACAGCATTTCT-----TGACCTC 1301

QY      268 AserSerLeuTrpLeuMetProPheSerLysAsnCysLeuGly-AspAsnThrSerTrpT 288
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1302 GAACATCTCTCG-----CGTTGTTTCGGCTGACGACCATCTGTGCTGCT 1343

QY      288 hrHisPheSerGlyAlaAlaLysLysAspLysArgValTyrHisCysArgGluGluG 308
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1344 TCCAT-----AACAAAT 1355

QY      308 IuTyrGlnSerLysTyrSerGlyTrpProThrAsnIleIleAsnIleGluIleProSer- 327
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1356 ATTACCAAGATATTACGAC-----ACTGCTGAACAATTAAATGTGAGCTATCCCGAAT 1409

QY      328 -----ThrSerValIThrGluThrAlaLeuValAsnMetThr 340
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1410 GGCTGAGCCCTGAAGAGACCTGAACCTTGTTAACAGACACGCGC-----A 1454

QY      340 erMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyT 360
```

```
Db      1455 AGGCGCTGGCAATGTGGCCACGCTGCTGGACAGCTGCTGTATAGCTTGCAGAGAA 1514
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      360 hrLeuPheLysAlaIleGluAspTyrAspProAspSer---LysArgMetLeuHisGln 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1515 CCACCTTACGCGACACAGATTAAGGCTGATCCCAACGGGTACCGCGCTGTATGGGT 1574

QY      379 euLysLysLeuTyrHisAspAspProValPheAsnProLeuThrProTTrpGluArgPro 399
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1575 TC-----CTGATTAAAGCCAAACATTCATGTTCCAGCTTA 1610

QY      399 roLeuLysAsnValPheCysIleTyrGlyAlaHisLeuLysThrGluValGly----- 416
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1611 TCTTCAGGACG-----GACCTTAAGCTCTTACCTGGGACGGCG 1649

QY      417 -----TyrTyrPheAlaProSerGlyLysProTyrProAspAsnITrIleIThr 434
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1650 CTCTTCAACATTACATGCTGCTGTCCAGCCC-----ACCA 1685

QY      434 epIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyIThrValAlaAspGlyA 454
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1686 ACACCACTTATGTTGTACAGATAGCTTGGCAAAATTTGACAGACAGTGTCAACCTCA 1745

QY      454 snAlaGlyProIleThrGlyAspGluThrValPro-----TyrHisSerL 469
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1746 CCCGAGACAGTGCACGATCCAAAGTAAAGTCCCAAGTAAAGAAAGATCTGTATGAGT 1805

QY      469 euSerTrpCysLysAsnITrPLeuGlyPro 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1806 ACTCATGGGTCCAG-----GGCCCT 1825
```

```
RESULT 3
US-10-131-826A-367
; Sequence 367, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OR INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131, 826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
```



```

; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO: 367
; LENGTH: 2906
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-367

Alignment Scores:
Pred. No.: 714 Length: 2906
Score: 85.00 Matches: 88
Percent Similarity: 31.21% Conservative: 44
Best Local Similarity: 20.80% Mismatches: 126
Query Match: 2.87% Indels: 166
DB: 1 Gaps: 22

US-09-651-651-5 (1-546) x US-10-131-826A-367 (1-2906)
QY 106 CysLysSerArgProkapsSerGlyLeuSerAlaIleThrGluLeuAap----- 121
DB 1430 TGCACACCTTCGGGAATCCCTAACCCTACACCGCCTATAAATAGATGAGCTGATCTT 1489
QY 122 -----ProGlyTyrIleThrGlyProLeuSerThrVal 132
DB 1490 TCTGGAACTATTATCTGCCATCAAGCCTGGCTTTCCAGGGT---TTGATGACACTT 1546
QY 133 TrrpLysGluTrrp----- 136
DB 1547 CAAAAACTGTGATGATACAGTCCAGATTCAAGTATGAAACGGAATCCTTGACAAC 1606
QY 137 LeuLysTrrpCysValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAap 156
DB 1607 CTTCAGTCACTAGTGAAGATCAACCTGGACACAAATATCTAATCTACTGCTCATGAC 1666
QY 157 TrrpArgLeuSerProThrLysLeuGluArgAapLeuTyrPheHisLysLeuLysLeu 176
DB 1667 ---CTCTTCACCTCCCTTGGATCATCTAGAGCGG---ATACATTTCAT----- 1708
QY 177 ThrPheGluThrAlaLeuLysLeuAargGlyProSerIleValPheAlaHisSerMet 196
DB 1709 -----CACAAACCT 1717
QY 197 GlyAasn---AsnValPheArgTyrPheLeuGluTrrpLeuAargLeuGluIleAlaProLys 215
DB 1718 TGGAACTGTAACCTGACATACATGCTGCTGAGTGTGATAAAGACATGGCCCCCTCG 1777
QY 216 His-----TyrLeuLys 219
DB 1778 AACACAGCTTGTGTGTCGCCGTGTAACTCTCCCAATCTAAGGGAGATGATGGA 1837
QY 220 TrrpLeuAargGluHisIleHisAlaTyrPheAlaValGlyAlaProLeuLeuGlySerVal 239
DB 1838 GAGCTGCACCAAGAT-----TACTTCACATGCTATGCTCGGTGATGTGGAGCCC 1888
QY 240 GluAlaIleLysSerThrLysSerGlyValThrPheGlyLeuProValSerGluGlyThr 259
DB 1889 CCTGCA-----GACCTCAATGCTCACTGAAGCAG 1918
QY 260 AlaArgLeuLeu-----SerAasnSerPheAlaSerSerLeuTrrpLeuMetPro 275
DB 1919 GCAGCTGAGCTGAATGCGGCGCTGCACATCCCTGATCTGTATCTTGATTTACTCCA 1978
QY 276 PheSerLysAasnCysLysGlyAapAanThrSerTrrpHisPheSerGlyGlyAlaAla 295
DB 1979 -----AATGAAACAGTATGACACAT-----GGGGCGTAC 2008
QY 296 LysLysAapLysArgValTyrHisCysAapGluGluGluTyrGlnSerLysTyrSerGly 315

```

```

DB 2009 AAGTGGCGGATAGCTGT-----CTCAGTGAT 2035
QY 316 TrrpProThrAnIleLeuAsnIleGluIleProSerThr----- 328
DB 2036 GGTACGTTAAATTTTCACAAATGTAACGTGTGAAGATCAGGCATGATACATGATGTGTG 2095
QY 329 -----SerValThrGluThrAlaLeuValAanMetThrSerMetClyCys 343
DB 2096 AGTAAATTCCTGGGAATACATCTGCTTCACACCAAGACCAAGTATGTAACCGCTTCAGCA 2155
QY 344 GlyLeuProThrLysLeuSerPheThrAlaArgLysLeuAlaAapGlyThrLeu----- 361
DB 2156 ---ACTCCTTT-CTCTTACTTTTCAACCGTCAAGTAGAGACTATGGAACCGTTCAGCA 2211
QY 362 -----PheLysAlaIleGluAapTyr 368
DB 2212 TGAGCAGCAGCACACAGATTAACAATGGGGTCCCACTCCAGTGTGCTGAGAGAACAC 2271
QY 369 AapProAap-----SerLysArgMetLeuHisGluLeu-----LysLysLeuTyrHis 384
DB 2272 CAATGTGACCACTCTCTTCACACCAAGACGACAAAGTGCAGAGAAACCTTCACCAT 2331
QY 385 AapAapProValPheAasnProLeuThrPro----- 394
DB 2332 CCCAGTGACTATATTAACAGTGGGATCCAGAAATTGATGATGATGATGAAGACTACCA 2391
QY 395 -----TrrpGluArgProProIleLysAasnValPheCysIleTyrGlyAlaHisLeu 411
DB 2392 AATCATCATTTGG-----GGTTTGGGCCATCACATCATGCG 2430
QY 412 LysThrGluValGlyTyrTyrPhe-----AlaProSerGlyLysProTyr 426
DB 2431 TGCACTGATGCTGGTGCATTTTCTACAAGATGAGAGAGACGACCATCGCAAAACCATCA 2490
QY 427 ProAapAasn 429
DB 2491 CGCCCCAAC 2499

RESULT 4
US-10-949-720-391
; Sequence 391, Application US/10949720
; Publication No. US20050249736A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Kerez, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITTING
; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; CURRENT FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 391
; LENGTH: 26000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-949-720-391

Alignment Scores:
Pred. No.: 2,266+04 Length: 26000
Score: 83.50 Matches: 84
Percent Similarity: 34.07% Conservative: 40
Best Local Similarity: 23.08% Mismatches: 132
Query Match: 2.82% Indels: 108

```



```

QY      238 ----- 238
Db      2746 GCTAATCAGCAGCAAAAACTGATGACGCTTCAATGAGGCAATTATGATATAT 2805
QY      239 VALGVALAIIeYserThrLeuSerGlyVal-----ThrPhe 251
Db      2806 ATTAATGCTGTTGTCATAGTGTGCTGGAGTACGATCGTACGCTTATATATACCTAT 2865
QY      252 GlyLeuProValSerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeu 271
Db      2866 TTGCTGATGATGATATACAGGTTTCTGCCGATGTGATCATCTTACAGTATTCAGAAAGCTATC 2925
QY      272 TrpLeuMetProPheSerLeuSerGlyValGlyAspAsnThrSerTrpThrHisPheSer 291
Db      2926 GCCCGATTTCACCTGATGCTTAAAC----- 2949
QY      292 GlyGlyAlaAlaLeuValSerLeuValSer-----ArgValTyr 302
Db      2950 ---CGGGCTTTAAACCGAGATGAGGTCAGCTTCGACCTGACCTTACCTCAGCTTC 3006
QY      303 HisCyAspArgGluGluGlyThrGlnSerLeuTyrSerGlyTyrProThrAsnIleIleAsn 322
Db      3007 TTCACGACTGCGGAAAGTATTAACAATTAACGTTACAGTACTGGGCTGTGCTCTGAAGTC 3066
QY      323 IIEGLUIEProSerThrSerValThrGluThrAlaLeuValAsnMetThrSerMetGlu 342
Db      3067 GTCCTATTATTCAGAAAACCTATGTGTGATCCCACTCAGCGCATTGGCGAAACCAAAATGATG 3126
QY      343 CyAGLYLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPhe 362
Db      3127 GATGGCGCTG-----TTGCATTCATCAACCAAGCAGGCAATTAATGCGATACGGTGAA 3180
QY      363 LySAlaIIeGluAspTyrArgProAspSerLeuArgMetLeuHis---GlnLeuValVal 381
Db      3181 GATGCTTTCAAAAACCTATTGTGACCAAGCTTGGACAGGTAGCAAACTGAAAGTAAATGAT 3240
QY      382 LeuTyrHisAspAsp 386
Db      3241 GCTTACCACGATATAT 3255

```

```

RESULT 6
US-11-147-725-3
; Sequence 3, Application US/11147725
; Publication No. US20050244878A1
; GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuer, Frits
; APPLICANT: Ward, Michael
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: BGJ5 Beta-Glucosidase and Nucleic Acids
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC697
; CURRENT APPLICATION NUMBER: US/11/147,725
; CURRENT FILING DATE: 2005-06-07
; PRIOR APPLICATION NUMBER: US 10/026,140
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-11-147-725-3

```

```

Alignment Scores:
Pred. No.: 743 Length: 1455
Score: 80.50 Matches: 77
Percent Similarity: 35.85% Conservative: 56
Beet Local Similarity: 20.75% Mismatches: 155
Query Match: 2.72% Indels: 83
DB: 8 Gaps: 20

```

```

US-09-651-651-5 (1-546) x US-11-147-725-3 (1-1455)
QY      155 TyrAspTrpArgLeuSerProThrIleuGluGluArgAsp-----LeuTyrPhe 171
Db      367 TACCACTGGAGATCTG---CTCAGGCGCTTACAGATCTCTATGAGGCTGGCTCAACGTG 423
QY      172 HisIleuValLeuThrPheGluThrAlaLeuValLeuValGlyGlyProSerIleVal 191
Db      424 GAAGAGCTCCAGCTGAGCTTTGAGGCGTATGGAGCTTG----- 462
QY      192 PheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeuGlyTrpLeuArgLeuGlu 211
Db      463 TCGTTGAACGTTTGGGAGCCGAGCTC-----CAGAACGTGATCAACATCAAC 510
QY      212 IleAlaProIleHisTyrLeuLeuThrPheLeuSerGlnIleHisAlaTyrPheAlaVal 231
Db      511 -----GAMCCCTGGAT---CAGGCCATCTATGATAT---GCCACC 546
QY      232 GlyAlaProLeuLeuGlySerValGluAlaIleYserThrLeuSerGlyValThrPhe 251
Db      547 GGCACCAAGCCCCCGGCGAGAGCAAGCATTAACAAGCACTCCACGAGGGCAACCTGCC 606
QY      252 GlyLeuProValSerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeu 271
Db      607 ACTGAGCGGCTGCGCTGGAAAGGCCCAAGATCATAGACCAATGCCGCGCTGCCCTGC 666
QY      272 TrpLeuMetProPheSerLeuSerGlyValGlyAsp-----AsnThrSer 286
Db      667 TACAGCAGGAGATTTGCGCCCTCGCAAAAGGCGCAATCGGATTCGCTCAACGGCGAC 726
QY      287 TrpThrHisPheSerGlyGlyAlaLeuValSerLeuValArgValTyrHisCyAspArgGlu 306
Db      727 TACTATGAGCCCTGGGAGACGAATGAGCTCGGCAAGAGAGCT-----GCT 774
QY      307 GlyGluGlyThrGlnSerLeuTyrSerGlyTyrProThrAsnIleIleAsnIleGlu----- 324
Db      775 GAGCAGCGAGTGAATTCACATTCGTGCTGTGGCAATCCCATCTCTTGAAAGAGAC 834
QY      325 -----IleProSerThrSerValThrGlu 332
Db      835 TATCCAGAGAGCATGAAGAGAGCTGGGCGGACGCTTCCAGCCCTCACTCCGCGGAC 894
QY      333 ThrAlaLeuValAsnMetThrSerMetGlu---CyAGLYLeuProThrLeuLeuSerPhe 351
Db      895 TTGGCCATCTCCATATGCGGAGAGACCGACCTTCAAGGATGATATTAACATCCGAG 954
QY      352 ThrAlaArgGluLeuAlaAspGlyThrLeuPheValIleGluAspTyrArgProAsp 371
Db      955 TTCGCGCGCCACCTA---GACGCTCCGCTC-----CCCGAGACGCGAC 993
QY      372 SerLeuValArgMetLeuHisGlnLeuValValLeuTyrHisAspArgProValPheAsnPro 391
Db      994 TATCTCGGCGCATCATGAGCAACGAGAAATTAAGACGAGGACCCGCTGGCGAGAG 1053
QY      392 Leu---ThrProTyrGluArg-----ProPheIle---LysAsnValPheCyAla 406
Db      1054 AGCGGCTCGGCTGGTGGCTCGCGCGGACATGTTCCGGAAGCATTCGCCGGGGTG 1113
QY      407 TyrGlyAlaHisLeuLeuThrGluValGlyTyrTyrPheAlaProSerGlyValProTyr 426
Db      1114 TACGGCTGTACGGAGGCC-----ATCTACATCAACGAGAGGAGATGCCGTGC 1164
QY      427 ProAspAsnTrpIleIleThr----- 433
Db      1165 CCTGGAGAGAGAAACATGACGCGGAGAGGCGGCTCAACGACCCCTTCGCGATCCGCTAC 1224
QY      434 ---AspIleIleTyrGluThrGluGlySerLeuValSerArgSerGlyThrValAlaAsp 452
Db      1225 TTGAGCTGCGCATTTGAGATCGATTTCCAAAGCCATTAACCAAGAGCGGCGTGGCAAG 1284
QY      453 GlyAsnAlaGlyProIleThrGlyAspArgIleValProTyrHisSerLeuSerTrpCyAs 472
Db      1285 GGGTACTTTGGCTGGCGCTTGGCTGAT-----AACTTGAATGTGTCA 1326

```

```

Qy      473 LysAsnTrpLeuGlyProLysValAsnIleThr 483
Db      1327 GATGGCTAC---GGACCCAGATTCCGGGCTCAGC 1356

RESULT 7
US-11-147-725-1
; Sequence 1, Application US/11147725
; Publication No. US20050244878A1
GENERAL INFORMATION:
; APPLICANT: Dunn-Coleman, Nigel
; APPLICANT: Goedegebuer, Frits
; APPLICANT: Ward, Michael
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: BGL5 Beta-Glucosidase and Nucleic Acids
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: GC697
; CURRENT APPLICATION NUMBER: US/11/147,725
; CURRENT FILING DATE: 2005-06-07
; PRIOR APPLICATION NUMBER: US 10/026,140
; PRIOR FILING DATE: 2001-12-18
; NUMBER OF SEQ. ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1991
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-11-147-725-1

Alignment Scores:
Pred. No.:      1.16e+03      Length:      1991
Score:          80.50         Matches:      77
Percent Similarity: 35.85%     Conservative: 56
Best Local Similarity: 20.75%   Mismatches: 155
Query Match:    2.72%         Indels:      83
DB:             8             Gaps:         20

US-09-651-651-5 (1-546) x US-11-147-725-1 (1-1991)
Qy      155 TyrAspTrpArgLeuSerProThrLysLeuGluIuArgAsp-----LeuTyrPhe 171
Db      428 TACCACCTGGATCTG---CCACAGCGCTTCACATGCTATGAGAGCGCTGCTAACGTG 484
Qy      172 HisLeuLeuLysLeuThrPheGluThrAlaLeuLysLeuArgGlyGlyProSerIleVal 191
Db      485 GAAGAGGTCCACGCTGACCTTGGACCGGATGCGAGGTTG-----523
Qy      192 PheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeuGluTrpLeuArgLeuGlu 211
Db      524 TGCTTTGAACGTTTTGGGACCGGAGTC-----CAGAACTGGATCACCATCAAC 571
Qy      212 IleAlaProLysHisTyrLeuLysTrpLeuAspGlnHisIleHisAlaTyrPheAlaVal 231
Db      572 -----GAWCCCTGGATT---CAGGCCACTATGAGATAT---GCCACC 607
Qy      232 GlyAlaProLeuLeuGlySerValGluAlaIleLysSerThrLysSerGlyValThrPhe 251
Db      608 GCGAGCAAGCGCCCGGAGAGAGACAGCATTAACAAGCACTCCACGAGCGCAACACTGCC 667
Qy      252 GlyLeuProValSerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeu 271
Db      668 ACTAGAGCCGTGCTCGCTGGAAAGGCCCAATGACCATGACCCCGCGCGCTGCGCTC 727
Qy      272 TrpLeuMetProPheSerLysAsnCysLeuGlyAsp-----AsnThrSer 286
Db      728 TACAGCAGGAGACTTTCGCCCTCGCAAAAGGCGCAGATCGGCATCTCCCTCAAGCGGAC 787
Qy      287 TrpThrHisPheSerGlyGlyAlaAlaLysLysAspLysArgValTyrHisCysAspGlu 306
Db      788 TACTATGAGCCCTCGGACAGCAATGAGCTCGGAGCAAGGAGGCT-----GCT 835
Qy      307 GluGluTyrGlnSerLysTyrSerGlyTyrProThrAsnIleLeuAsnIleGlu----- 324

```

```

Db      836 GAGCGAGCATGGAAATTTTCACATTCGGCTGTGGCCATCCCATCTTTGAAGAAGAC 895
Qy      325 -----IleProSerThrSerValThrGlu 332
Db      896 TATCCAGAGAGCATGAAGAAGCAGCTGGCGAGAGGCTTCCAGCCCTCCTCCCGCGAC 955
Qy      333 ThrAlaLeuValAsnMetThrSerMetGlu---CysGlyLeuProThrIleLeuSerPhe 351
Db      956 TTTCGCATCTCCATATCCGAGAGACGACTTCTACGCGATTAATTAACATACATCCAG 1015
Qy      352 ThrAlaArgGluLeuValAspGlyThrLeuPheLysAlaIleGluAspTyrAspProAsp 371
Db      1016 TTTCGCGCGCACCTA---GACGCTCCCGTC-----CCCGAGAGGAC 1054
Qy      372 SerLysArgMetLeuHisGlnLeuLysLysLeuTyrHisAspAspProValPheAsnPro 391
Db      1055 TATCTCGCGCGCATTCATGACGACGAGAAATTAAGACGCGAGCCCGTTGGCGAGAG 1114
Qy      392 Leu---ThrProTrpLysArg-----ProPheIle-----LysAsnValPheCysIle 406
Db      1115 AGCGGCGCTGCGCTGCGCTGCTGCGGACATGTTCCGAAAGCATCTCGCCCGGATG 1174
Qy      407 TyrGlyAlaHisLeuLysThrArgLysValGlyTyrTyrPheAlaProSerGlyLysProTyr 426
Db      1175 TACGCGCTGACGGCAAGCCC-----ATCTACATGACCGAGAACGGAATGCCGTGC 1225
Qy      427 ProAspAsnTrpIleIleThr-----433
Db      1226 CCTGAGAGAGAGAAACATGACGTGCGAGAGCGCGTCACAGCCCTTCGCGATCCGATAC 1285
Qy      434 ---AspIleIleLysTrpLysArgLysSerLeuValSerArgSerGlyThrValValAsp 452
Db      1286 TTTCAGCTCGGACTTGGACCTGATTTCCAAAGGCCATTAACGAGACGCGCTGCTCAAG 1345
Qy      453 GlyAsnAlaGlyProIleThrGlyAspGluThrValProTyrHisSerLeuSerTyrCys 472
Db      1346 GGGTACTTTCGCGGCGGCTGCTCGAT-----AACTTGAATGCTCA 1387
Qy      473 LysAsnTrpLeuGlyProLysValAsnIleThr 483
Db      1388 GATGGCTAC---GGACCCAGATTCCGGGCTCAGC 1417

RESULT 8
US-10-957-569-11
; Sequence 11, Application US/10957569
; Publication No. US20050246785A1
GENERAL INFORMATION:
; APPLICANT: COOK, Zhifeng et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11
; LENGTH: 2310
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-957-569-11

Alignment Scores:
Pred. No.:      2.01e+03      Length:      2310
Score:          79.00         Matches:      56
Percent Similarity: 33.82%     Conservative: 36
Best Local Similarity: 20.59%   Mismatches: 100
Query Match:    2.66%         Indels:      80
DB:             1             Gaps:         13

US-09-651-651-5 (1-546) x US-10-957-569-11 (1-2310)

```

```

Oy      146  GluAlaAsnAlaIleValAlaValProTyrAsp----- 156
Db      1029  GAATTCGGGTGAACCTCTTCGGCTTTAAGCTACGACGTCGTTTCAAGCCTTACCTAAATAC 1088
Oy      157  TrpTArgLeuSerProThrLysLeuGluGluArgAspLeu----- 169
Db      1089  TTCGATTCTCAACCGGACGGAACTAAATCACCAGACGTCGAGATTCAAGCTGATCAGCA 1148
Oy      170  ---TyrPheH1sLysLeuLysLeuThr----- 177
Db      1149  ACGATGATGACAGCATTTGCGATTACAGAGAACTTCGTCGTCGTAACCTGACCAAGATC 1208
Oy      178  ---PheGluThrAlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHisSerMet 196
Db      1209  GTTTTCAAGCTCCCGGAGATATCCGCGGTGGTCTCCGGGTGTTTACAGAACG----- 1262
Oy      197  GlyAsnAsnValPheArgTyrPheLeuGluTrpLeuArgLeuGluIleAlaProLysHis 216
Db      1263  ---AACAAAGTCGCAGAGATTCGGGATT-----TTAGCAAAATACCGCGGAAGATTCAACG 1313
Oy      217  TyrLeuLysTrpLeuAspGlnHis-----IleH1sAlaTyrPheAlaValGly 232
Db      1314  AACATTAAAGTGAATTGATGATGCTCAGATTGCTTCGCTTCATCTGTGAACGCTGGGAA 1373
Oy      233  AlaPro-----LeuLeuGlySerVal----- 239
Db      1374  GAGCGAGAAACAGATGAAGTCGCTGTATAGGCTCTGTATGATCCACAGACTCAATT 1433
Oy      240  -----GluAlaIleLysSerThrLysSerGlyValThrPheGlyLeuPro 254
Db      1434  TTCAACGAGCTCGACGAGATCTCAAGAGTGTCTGTCTGAAATCCGCTGAATTCAAA 1493
Oy      255  ValSerGluGlyThrAlaArg---LeuLeuSerAsnSerPheAlaSer---SerLeuTrp 272
Db      1494  ACCGGTGAATCAACTCGCCGTCGATCATCTCCAAGAAAGATCAACAAGTCAACTCGAA 1553
Oy      273  LeuMetProPheSerLysAsnCysAllyseGlyAspAsnThrSerTrpThrHisPheSer--- 291
Db      1554  GCAGGATGCTCAACAGAAACATGCTCGGCGCTTAACCAATATGCTTACTTGCTTTA 1613
Oy      292  -----GlyAlaIleAlaLysLysAspLysArgValTyrHisCys 304
Db      1614  GCCGAGCCGTGGCCTTAAAGTCTCAGAGATTCCGCTAAAGTTGATCTCACTACGAGAACTT 1673
Oy      305  AspGluGluGluTyr---GlnSerLysTyrSerGlyTyrProThrAsnIleIleAsn--- 322
Db      1674  AAGAAACATCTTAAAGCGCATTAACCGTTACGAGAGAGCCCTGTTTCTCCCGGAGAA 1733
Oy      323  -----IleGluIleProSerHisSerVal 330
Db      1734  CGAGGAGAGAGAACGAAAGATATACCTCTGTTCTGTACAGACGAGAACAGATGAAA 1793
Oy      331  ThrGluThrAlaLeuValAsnMetThrSerMetGlu 342
Db      1794  TCGAGATTACAGATGATTAAACCCGTTACTTACGAG 1829

RESULT 9
US-10-467-962B-78
/ Sequence 78, Application US/10467962B
/ Publication No. US20050246784A1
/ GENERAL INFORMATION:
/ APPLICANT: Plesch, Gunnar
/ APPLICANT: Blau, Astrid
/ APPLICANT: Daeschner, Klaus
/ APPLICANT: Klein, Mathieu
/ TITLE OF INVENTION: Identification of Herbicidally Active Substances
/ FILE REFERENCE: 2000_857
/ CURRENT APPLICATION NUMBER: US/10/467,962B
/ CURRENT FILING DATE: 2003-08-14
/ PRIOR APPLICATION NUMBER: PCT/EP02/01466
/ PRIOR FILING DATE: 2002-02-13
/ NUMBER OF SEQ ID NOS: 109
/ SOFTWARE: Seq ID NOS: 2.0

```

```

; SEQ_ID NO 78
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1023)
; OTHER INFORMATION: ORF F25123_240, Linle 126905
US-10-467-962B-78

Alignment Scores:
Pred. No.: 709 Length: 1023
Score: 78.50 Matches: 73
Percent Similarity: 29.73% Conservative: 37
Best Local Similarity: 19.73% Miscellaneous: 111
Query Match: 2.65% Indels: 149
DB: 1 Gaps: 16

US-09-651-651-5 (1-546) x US-10-467-962B-78 (1-1023)
Qy 126 ThrGlyProLeuSer-----ThrValTrpLysGluTrpLeuLysTrpCyValGluPhe 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 ACCGCCCACTGAGCCAAAGATGGAGTGGTCAGCGTGGTCCATTGACTGACGAGCAT 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 144 GylIleGluAlaAspAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLys 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 GGT-----CCGAATCCAGTGGTGGCAATTGCCCTACAGAAAGAGTTCGCCGAGCATTA 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 164 LeuGluGluArgAspLeuTyrPhe-----HisLysLeuLysLeuThr 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 GATTACTTCGGTGGCATTTACTTTTCCGACGAGCATCTCCCGCGCACTACGACTCAG 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 178 PheGluTrpAlaLeuLysLeuArgGlyGlyProSerIleValPhe----- 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 GAAGAAACCCCTCTC---TTAAATCCCGCACTACACAGTGGCATTTTCAGCGCCTA 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 192 ----- 192
Db 289 GTACTCGAGGCCCTTAATACGACTGTGTTGAAGAACTCGACTTCATCGAACGCAATGCT 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 193 ----- 196
Db 349 GAGGACTACTCTTAAGAACTACCAACTGTGGCATCTACGGCGCATGGGTTCCAGAGAACTG 408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 197 GlyAsnAsnValPheArgTyrPheLeuGluTrpLeuArg-----LeuGluIleAlaPro 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GGTCCCTGATGTTGCAGAGGAGAGAACTTGAAATTTACCCGTAGAGTACTTTCACCTTGATGCC 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 215 LysHisTyrLeuLysTrp----- 220
Db 469 AAACATTATTCATGCTTGTCACATGAGCAGTGGACCTACGGGCAATTAGAGAGATGGAA 528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 221 -----LeuAspGlnHisIleHisAla----- 227
Db 529 GATGAGCTCGAT---TACTGTACACAGCTCTTGAAGCTGACGCTTTTAACAATTCGCC 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 228 -----TyrPheAlaValGlyAlaProLeuLeuGlySerValGluAlaIle 242
Db 586 TGGATTCAGAGGATTAATGATCATCAACCAATCTCCCTTGTGGAGAGCCTAGAAAGCATG 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 243 LysSerThrLeuSerGlyValThrPheGlyLeuProValSerGluGlyThrAlaArgLeu 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 AGAGATCT-----GAAGTAAAGCTTACCAATCAAAAGCCATT 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 263 LeuSerAsnSerPheAlaSerSerLeuTrpLeuMetProPheSerLysAsnGly 282
Db 682 TTAAACCAAT-----CTGCA 696
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 283 AspAsnThrSerTrpThrHisPheSerGlyGlyAlaAlaLysLysAspLysArgValTyr 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 AACGAGAGCTCATGGGATACCTA-----AAAGCTCTTTAC 732
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 303 HisCysAspGluGluGluTyrGlnSerLysTyrSerGlyTrpProThrAsnIleLeuAsn 322

```

```
Db 733 AAAGACGACAAAGATCTCG----- 753
Qy 323 ILGLIleProSerThrsValThrgluThralAleuValAsnMetThrsMetGlu 342
Db 754 -----ATTAGTGATGCCAAGTGTTCCTCGAGCTGTTGAATGTTCTATCCCGACAGAT 807
Qy 343 Cys-----GlyLeuProThrLeuLeuSerPheThralArgGluLeuValAsp 358
Db 808 TGCTTCATGATGATTCGCTCGAGACACCTTTTGGATCTT-----CTATGATGAT 855
Qy 359 Gly-----ThrLeuPheLysAlaIleGluAspTyrAspPro 370
Db 856 GGACTGAGACCAACCAACGACATPAAAGCTCACTGAGCTCTPAGCTAATGAAGAACCA 915
Qy 371 AspSerLysArgMetLeuHisGlnLeuLysLeuTyrHisAspAspProValPheAsn 390
Db 916 GAGACTAAGCTTGGCCAAATTTGGTGTGTAATTTCTGCTGTGATGATCCCTGTAAGAGCT 975
Qy 391 ProLeuThrProTyrGluArgProProIle 400
Db 976 AACTATTGGGCGATGAGAGAAAGACAGATT 1005

RESULT 10
US-10-518-599-3
; Sequence 3, Application US/10518599
; Publication No. US20050251873A1
; GENERAL INFORMATION:
; APPLICANT: PENNINGER, JOSEPH M.
; APPLICANT: CRACKOWER, MICHAEL A.
; TITLE OF INVENTION: ACE2 ACTIVATION FOR TREATMENT OF HEART, LUNG AND
; FILE REFERENCE: SONN: 064US
; CURRENT APPLICATION NUMBER: US/10/518, 599
; PRIOR FILING DATE: 2004-12-17
; PRIOR APPLICATION NUMBER: PCT/CA03/00882
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/389, 709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2739
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-518-599-3

Alignment Scores:
Pred. NO.: 3.22e+03 Length: 2739
Score: 78.00 Matches: 61
Percent Similarity: 31.02% Conservative: 42
Best Local Similarity: 18.37% Mismatches: 121
Query Match: 2.63% Indels: 108
DB: 1 Gaps: 12

US-09-651-651-5 (1-546) x US-10-518-599-3 (1-2739)
Qy 245 ThrLeuSerGlyValThrPheGlyLeuProValSerGluGlyThrAlaArgLeuLeuSer 264
Db 203 AGCCTTGCTGGCTGTACTACTGCTCAGTCCCTCACCGAGAAATGCCAAGACATTTTTA 262
Qy 265 AsnSerPheAlaSerSerLeuTriPheMetProPheSerLysAsnCysLysGlyAspAsn 284
Db 263 AACCAACTTT-----AATCAGGAAGCTGAGAGAC 289
Qy 285 ThrSerTyrThrHisPheSerGlyGlyAlaAlaLysLysAspLysArgValTyrHisCys 304
Db 290 CTGTCT----- 295
Qy 305 AspGluGluGluTyrGlnSerLysTyrSerGlyTyrProThrAsnIleIleAsnIleGlu 324
Db 296 -----TATCAAGTTCACCTTGCTTCTTGGAATTTAATTAATCAATTAAGTAA 343
```

```
Qy 325 IleProSerThrsValThrgluThralAleuValAsnMetThrsMetGluCysGly 344
Db 344 GAAATGCCCAAAAGATGAGTGTGCGAGCCAAATGCTTGCTTTATATGAACAG 403
Qy 345 LeuProThrLeuLeuSerPheThralArgGluLeuValAspGlyThrLeuPheLysAla 364
Db 404 TCTTAAGCTGCCCAAGATTCTTCATCAAGAAATCACAAGATCCGATCATCAAGCGTCA 463
Qy 365 IleGlu-----AspTyrAspProAspSer 372
Db 464 CTACAGACCCCTTCAGCAAGATGGGTCTTCAGACCTTCAGACAGCAAGAACAGTTG 523
Qy 373 LysArgMetLeuHisGlnLeuLysLysLeuTyrHisAspAspProValPheAsnProLeu 392
Db 524 AACCAATTCGTGAACACATGAGCACCATTTTACGTACCTGAAAGATTGCAACCCAAAG 583
Qy 393 ThrProTyrGlu-----ArgProProIleLysAsnValPheCysIle----- 406
Db 584 AACCCCAAGAAATGCTTATTACTTGAAGCCAGATTTGATGAATAATATGGCGACAGACA 643
Qy 407 ---TyrGlyAlaHisLeu-----LysThrgluValGlyTyrTyrPhe 419
Db 644 GACTAAGACTGTAGGCTGTGGCATGAGAGGCTGAGAGGCTGAGAGGCTGAGGTTGGCAAGCATG 703
Qy 420 AlaPro----- 421
Db 704 AGCCGTTGTATGAGATGATGTGCTCTGAAAAAGAGATGCAAGCAACAAATTAT 763
Qy 422 -----SerGlyLysPro 425
Db 764 AACGATATTGGGATATTGGAAGGGAGCTATGAACAGAGGAGGACAGATGGCTCAAC 823
Qy 426 TyrProAspAsnTriPheIleThrAspIle-----Ile 436
Db 824 TATTAACCGTAAACAGATGATGAAGATGTGAACGTACCTTCGACAGATCAAGCCATTTG 883
Qy 437 TyrGluThrGluGlySerLeuValSerArgSerGlyThrValValAspGlyAsnAlaGly 456
Db 884 TATGAGCATTTGATGCTTATGTGAGAGAGAG-----TTGATGATTAACCTTACC 937
Qy 457 ProIleThrglyAspGluThrValProTyrHisSerLeuSer-----TyrCysLysAsn 474
Db 938 TACATCAGCCCCACCTGATGCTCCTCCGCCCATTTGCTGTGATATGTGGGTAGATT 997
Qy 475 Tyr-----LeuGlyProLysValAsnIleThrMetAlaProGlnProGluHisAspGly 492
Db 998 TGGACAAATCTGTACCT---TTGACTGTTCCCTTGACAGAAACCAACATATAGATGTT 1054
Qy 493 SerAspValHisValGluLeuAsnValAspHisGlu 504
Db 1055 ACTGATGCAATGATGATGATCAGGGCTGGGATGCAGAA 1090

RESULT 11
US-11-042-988-10
; Sequence 10, Application US/11042988
; Publication No. US2005024481A1
; GENERAL INFORMATION:
; APPLICANT: SILICIANO, ROBERT
; APPLICANT: ZHANG, HAILI
; APPLICANT: ZHOU, YAN
; TITLE OF INVENTION: SINGLE CELL ANALYSIS OF HIV REPLICATION CAPACITY AND
; FILE REFERENCE: 62760(71699)
; CURRENT APPLICATION NUMBER: US/11/042,988
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/540,716
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 9719
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
```

US-11-042-988-10

Alignment Scores:

Pred. No.:	2.18e+04	Length:	9719
Score:	77.50	Matches:	75
Percent Similarity:	30.75%	Conservative:	52
Best Local Similarity:	18.16%	Mismatches:	133
Query Match:	2.61%	Indels:	153
DB:	9	Gaps:	18

US-09-651-651-5 (1-546) x US-11-042-988-10 (1-9719)

QY 45 llelleProGlyPheAlaSerThrGluLeuArgAlaTrpSerIleLeuAspCysProTyr 64

DB 1291 GTGATACCCATGTTTTCAGCATTTATCAGAAAGAGCC----- 1326

QY 65 ThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrTrpLeuLeuSer 84

DB 1327 ACCCCACAAAGATTTTAAACCCATG-----CTAAACACAGTGGGGGACATCAA 1374

QY 85 AlaValAsnCysTrpPheLysCysMetValLeuAspProTyrAsnGlnThrAsp----- 102

DB 1375 GGAGCCATGCMAATGTTAAAGAGACCATCATGAGAAAGCTGCAGAAATGGGATGAGGTG 1434

QY 103 HisPro-----GluCysLysSerArgPro 110

DB 1435 CATCCAGTGCATGACAGGCTTATTCACAGCCAGCCAGATGAGAAACCAAGGGAAATGAC 1494

QY 111 AspSerGlyLeuSerAlaIleThrGluLeuAspProGlyTyrIleThrGly----- 127

DB 1495 ATGACAGAACTACTAGTACCTTCAGAAACAAATAGATGATGACAAATATCCACCT 1554

QY 128 ---ProLeuSerThrValTrpLysGluTrpLeuLysTrpCysValGluPheGlyIleGlu 146

DB 1555 ATCCCAAGTGGAGAAATTTTAAAGATGATGATATC-----CTGGGA 1596

QY 147 AlaAsnAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuGlu 166

DB 1597 TTTAAATTAATTAATTAAGAAATG-----TATAGCCCTTACAGCATT----- 1635

QY 167 ArgAspLeuThrPheHisLysLeuLysLeuThrPheGluThrAlaLeuLysLeuArgGly 186

DB 1636 -----CTGGACATTAAGACAA 1650

QY 187 GlyProSerIleValPheAlaHisSerMetGlyAsnAsnValPheArgTyrPheLeuGlu 206

DB 1651 GGACCAAGGAAACCTTTAGACATATGTAGACCG-----TTCTATAA 1695

QY 207 TrpLeuArgLeuGluIleAlaProLysHisTyrLeuLysTrp----- 220

DB 1696 ACTCTAAGAGCCAGCAAGCTTCAACAGAGGTAAAAATTGATGACAGAAACCTTGTTG 1755

QY 220 ----- 220

DB 1756 GTCCAAATGCGAACCCAGATGTGAAGATTTTAAAGCATTTGGGACAGCGGTACA 1815

QY 221 LeuAspGlnHisIleHisAlaTyrPheAlaValGlyAlaPheLeuGly---SerVal 239

DB 1816 CTAGAAGAAATGTATGACAGATGTACAGAGTAGAGAGACCC-----GGCATTAAGCA 1869

QY 240 GluAlaIleLysSerThrLeuSerGlyValThrPheGlyLeuProValSerGluGlyThr 259

DB 1870 AGAGTTTGGCTGAAGCATGAGCCAAAGTAA-----AATTGACGTACC 1914

QY 260 AlaArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMetPro----- 275

DB 1915 ATATATGATGACAGAGCAATTTTACCAACCAAGAAAGATTTTAAAGTGTTCATTTG 1974

QY 276 -----PheSerLysAsnCysValGlyAspAsnHisSerTrpThrHisPhe 290

DB 1975 GGCAAAAGAGGACACACAGCCAGAAATTTGACG----- 2007

QY 291 SerGlyValAlaValLysLysAspLysArgValTyrHisCysAspGluGluTyrGln 310

DB 2008 -----GCCCTAAGAAAAAGGCTGTGGAAATGTGAAAGAAAGACACCA 2055

QY 311 SerLys-----TyrSer 314

DB 2056 ATGAAAGATTTGACTGAGAGACAGGCTAATTTTAAAGAAATCTGCTTCTTCAAG 2115

QY 315 GlyTrpProThrAsnIleIleAsnIleGluIleProSerThrSerValThrGluThrAla 334

DB 2116 GGAAGCCAGGAGATTTTCTTACAGACAGACCAAGCCAAAGCCCAACAGAA----- 2169

QY 335 LeuValAsnMetThrSerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArg 354

DB 2170 -----GAGAGCTTCAGTCTGGGGTAGACAAACAACTCCCTCAGAAAGCAG 2217

QY 355 GluLeuAlaAspGlyThrLeuPheLysValIleGluAspTyrAspProAspSerLysArg 374

DB 2218 GAGCCGATTAAGACAAAGAACTGTATCT----- 2244

QY 375 MetLeuHisGlnLeuLysLysLeuTyrHisAspAspPro 387

DB 2245 ---TTAACTTCCCTCAGGCTCTTTGGCAACGACCC 2280

RESULT 12

US-10-663-794-3/c

; Sequence 3, Application US/10663794

; Publication No. US20050244833A1

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CLO01164CIP-DIV II

; CURRENT APPLICATION NUMBER: US/10/663,794

; CURRENT FILING DATE: 2003-09-17

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FaastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 48763

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-663-794-3

Alignment Scores:

Pred. No.:	2.12e+05	Length:	48763
Score:	77.50	Matches:	98
Percent Similarity:	29.11%	Conservative:	63
Best Local Similarity:	17.72%	Mismatches:	165
Query Match:	2.61%	Indels:	230
DB:	1	Gaps:	22

US-09-651-651-5 (1-546) x US-10-663-794-3 (1-48763)

QY 130 SerThrValTrpLysGluTrpLeuLysTrpCysValGluPheGlyIleGluAlaAsnAla 149

DB 20296 TCTCAACGTATGAATGATGATTAATGAGACAGATGATTTG----- 20255

QY 150 IleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuGluArgAspLeu 169

DB 20254 -----GAACTTAACATG 20243

QY 170 TyrPheHisLysLeuLysLeuThrPhe-GluThrAlaLeuLysLeuArg----- 185

DB 20242 TTTATTAACACTGTATGATACATGATCTGAGACATCGCCTTAATGTAGTGACTGTAAATG 20183

QY 186 -----GlyLysProSerIleValPheAlaHis-Ser 196

DB 20182 ACCTGACGAAATGAAGGAGGACAGCAGTGTGACCCAAATAATTAATCTGCACGTAC 20123

QY 196 etGlyAsnAsnAlaPhe---ArgTyrPhe----- 204

DB 20122 AAGGAAACGATCATTAACACAGCATTTACACACCAAGGGCAATGCAATCTTATATCAA 20063



```

QY 205 --leuglurp-----leuargleugluile 213
    |||
    |||
    |||
Db 20062 AATTATCTCGAAATCCGGCAGAAAGATGCACTGAGACCAAGACTGCTTAA 20003
QY 213 laProlyshIeTYrleu-----lyeTTrpleu 222
    |||
    |||
    |||
Db 20002 ATCCCAACACACACTGGAATCTTCTGGAACTGGAAGGATGCCATTTCAATGGCTGG 19943
QY 222 spgIn-----HieIleHialaTyRhealValaIglValaProleuIeugIys 238
    |||
    |||
    |||
Db 19942 GCAGCAGATAAGTTAGAGACTACATTCAGA---GTCATATCAGAGCCCTGCCAAGTGGCT 19886
QY 238 er-----ValgluaIaIlyseTThreIeus 247
    |||
    |||
    |||
Db 19885 CACACATGAGACAGACTCGGTCATCAAGGAAGGACCATTTTCATACACAGTCAGTGA 19826
QY 247 erGlyValThrPheGlyLeuProValSer----- 256
    |||
    |||
    |||
Db 19825 GTGGCATTTTAATGAATTGGCCCAATGTCAATTTTTCATTTCTTACTTTTCTTTTAA 19766
QY 257 -----GlugIyThralaRgleuLeuSerAenSerPheAlaSerSerI 271
    |||
    |||
    |||
Db 19765 AGACACAAATGGCTGGGCGCAGTAGACTCAGCGCTGTATCCACAGATTTTGGAGGCGCA 19706
QY 271 eu----- 271
Db 19705 GCGGGTGATCACTGAGTCAAGAGTTCACACAAAGCTGACCAATGATGAAACC 19646
QY 272 -----TrpleuMeCProhes 277
    |||
    |||
    |||
Db 19645 CATCTCTACTAACATACAAATTAAGTCTGCACACAAAGCGGCTTTTATCTCACAGCTG 19586
QY 277 erIySaenCylyeGlyAspAenThrSerTThrHisPheSer---GlYglYAlaIal 296
    |||
    |||
    |||
Db 19585 CTCGGAGGCTGAGGAGAGAAAT-----CACTTGAACCTGGAGAGTGCAGCTT 19538
QY 296 yslYsaPlysaRgValTYrHisCyasaPglugIuTyrgInSerIySerGlyT 316
    |||
    |||
    |||
Db 19537 GTAAGTAGACAGAGATCGTACACAGCTGCACT-----TCAGCCT 19502
QY 316 rPProThraInIleIlaSnIleGluIleProSerThrSerValThrGluThraIleu 336
    |||
    |||
    |||
Db 19501 GGGTGACAGAGTAGAGACTCGCTC-----TCAAAAAAAGAAACCAATCTTGC 19451
QY 336 alAenMetThrSerMetGluCyseGly-----LeuProThrLeuLeuSerPheThra 353
    |||
    |||
    |||
Db 19450 TCTG-TGCCCCAGGCTGGAATGCAGTGTGAAATCAAGATCTAATCACTGCAGCCT 19392
QY 353 laarGluLeuIlaaPglYThrLeuPheIyalaIleGluAspTYrAspProAspSerI 373
    |||
    |||
    |||
Db 19391 CGAACTCCTGGACTCAAGGGAGCCTCACCCACTCATTTCTCAGAGTAGCTGAGACTATA 19332
QY 373 ysaIyMetIeuhIeGluLeuIylyIeIuTYrHisAspAspProValPheAsnProleuT 393
    |||
    |||
    |||
Db 19331 GGCC-----A 19326
QY 393 hrPProTPrGluArgProProIleIySaenValPheCyIleTYrGlyAlaHisIeIuYr 413
    |||
    |||
    |||
Db 19325 CACCAAC-AACTCGCTAATTTTAAATTTTGTTCAGAGTAGAGGCTCATATATGCA 19267
QY 413 hrGluValaIglYTYrTYrPheAlaProSerGlyIyysProTYrProAspAsnTrIleIeT 433
    |||
    |||
    |||
Db 19266 CCCAGGCTGGTCAAACTCTAGACTCAAGCAATCCT----- 19229
QY 433 hrAspIleIleTYrGluThrGluIySerIeIuValSerArgSerGlyThrValaIaSpG 453
    |||
    |||
    |||
Db 19228 -----CCCACTTGCTTCCCAAGCAGCTGGGATTCACAGCA 19192
QY 453 lyAaenIaIglYProIleThrGlyAspGluThrValProTYrHisSerIeu----- 469
    |||
    |||
    |||
Db 19191 TGAACATCGCACTCAGGCTGCACTGAATCACTCTTGCCACCCCTGCTTGTTC 19132
QY 470 --SerTPrCyalyaSnTrIleuGlyProIyValaSnIleThrMeCAla----- 485

```

```

Db 19131 ACAGGTGAGATTGAATCTCTGTGTACAGGCTTATATGATGCGCGGATTGAGAG 19072
QY 485 ----- 485
Db 19071 TCTTTCAGCACACAGCCCAACCCCTCACTTCAGCAATCATCTGAATACACTTCCTT 19012
QY 486 -----ProGInProGluHIsaPglYSerAspValHisVal----- 497
    |||
    |||
    |||
Db 19011 GGTCCCTTGCGGAACAGGCAAGAGGTGGTGGATCCACATGAGATGTCGACACTGT 18952
QY 498 --GluIeSaenValaPheHisGluHIsaPglYSerAspIleIleIlaaenMetThrIyAlaP 517
    |||
    |||
    |||
Db 18951 AAAAAACAATCTTAAGGCACTTACATGGGAA-----TGCCATGAAGAAACAAATTTAA 18898
QY 517 roArGValIySTYrIleThrPheTYrGluAspSer 528
    |||
    |||
    |||
Db 18897 AAGCTGCAGGTATTT-TCCTTTTACCCTGACAAAT 18864

```

```

RESULT 13
US-11-074-176-247
; Sequence 247, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kjaehammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Thereof
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 247
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1287)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0) ... (0)
; OTHER INFORMATION: ORF 1892; adenylosuccinate synthase
US-11-074-176-247

```

```

Alignment Scores:
Pred. No.: 1,38e+03 Length: 1287
Score: 77.00 Matches: 61
Percent Similarity: 34.08% Conservative: 45
Best Local Similarity: 19.61% Mismatches: 115
Query Match: 2.60% Indels: 90
DB: 7 Gaps: 16

```

```

US-09-651-651-5 (1-546) x US-11-074-176-247 (1-1287)
QY 145 lIeGluIlaaSnIaIleValaIaValaProTYrAspTTrPArg----- 158
    |||
    |||
    |||
Db 292 ATTTCTAACCGTCTCACATTATTAATGCCCTTACAGACTTAAGACAGATGAATACCAAGAA 351
QY 159 -----LeuSerProThrIyLeu 164
    |||
    |||
    |||
Db 352 GAAGCTAAGGTAAATAAAGATTGGTACTTAAGAATGTGATTTGGTCCAACTTACATG 411
QY 165 Glu-----GluArgAspLeuTYrPhe 171
    |||
    |||
    |||
Db 412 GATTAAGCTTCAAGAAATAGGTATTCCGGTTTGTGATTACTCGAAAAAAGATACCTTTGAA 471

```



```
/ CURRENT FILING DATE: 2005-03-22
/ PRIOR APPLICATION NUMBER: 60/555,720
/ PRIOR FILING DATE: 2004-03-23
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn Ver. 3.3
/ SEQ ID NO 1
/ LENGTH: 1824
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: nucleotide sequence
US-11-085-864-1
```

```
Alignment Scores:
Pred. No.: 2.85e+03 Length: 1824
Score: 76.00 Matches: 67
Percent Similarity: 32.81% Conservative: 38
Best Local Similarity: 20.94% Mismatches: 80
Query Match: 2.56% Indels: 136
DB: 9 Gaps: 23
```

US-09-651-651-5 (1-546) x US-11-085-864-1 (1-1824)

```
QY 54 leuargAlaTrpSerIleLeuAspCysProTyrThrProLeuAspPheAsnProLeu--- 72
DB 953 ATTAATGCGTGGTGC-----TGCACCATCAGCAGCTTATCG--AAATCCTTGCCA 906
QY 73 -----AspLeuValTrpLeuAsp 78
DB 905 CGTAAGTCCGCACTTCATGACGACCAAGCCAGTAAGTAGAACGGTTGTGGTGA--- 849
QY 79 ThrThrIysLeuSerAlaValAsnCyseTrpPheIysCysMetValLeuAspProTyr 98
DB 848 -----ATCAGAACTGTTGC-----CCCTTC 828
QY 99 AsnGlnThrAspHisProGluCysLysSer-----ArgProAspSerGlyLeuSer 115
DB 827 ACTGCCACTGACCGGATGCCGACCGAAGCGGTAGATATCATCAGACTCTGTCGGCTTTTG 768
QY 116 AlaIleThrGluLeu-----AspProGlyTyr----- 124
DB 767 GCTGTGACGACAGTTCATAGATAACTTACCCGGTGGCAGAGGTGCGAATTCACC 708
QY 124 ----- 124
DB 707 ACTTGCAAAAGTCCCGCTAGTGCCTTGTCCAGTTGCAACCACTGTGATCCGCATCAGC 648
QY 125 -----IleThrGlyProLeuSerThrValTrpLysGlu-----TrpLeuLys 138
DB 647 AGTTCACACGCTGATCATCATGTGGCCACCACTGGCAGTCAACAGACGCGGTACAG 588
QY 139 TrpCysValGlu-----PheGlyIleGluAla 147
DB 587 TCTTGGCGGACATGCGTACCAACGATATCGTCCACCAACGAGGTGTCGGGTG--- 534
QY 148 AsnAlaIleValAlaValProTyrAspTrpArgLeuSerProThrLysLeuGluLysArg 167
DB 533 -----GTGTAGAGCATTAAGCTGGATG-----ATTCGGCATAGTTAAAGAAATCA 486
QY 168 -----AspLeuTyrPheHisLysLeuLysLeuThrPheGluThrAlaLeuLysLeu 184
DB 485 TGAAGATGAAGACTGCTTTTTC-----TTGCGGTTTTCGTGCGTATATCACCATT 438
QY 185 ArgGlyGlyProSerIle-----ValPheAlaHisSer----- 195
DB 437 CCGCGCGGGGAT-AGCTCTGCCAGTTCAGTTTCGTTGTTACACAAACGCGTGAATCAGTACT 379
QY 196 ---MetGlyAsnAsnValPheArgTyrPhe---LeuGluTrpLeuArgLeuGluIleAla 213
DB 378 TTTCGGCGCATATACATACGCGGTGACATCGGCTTCAATAG---CGTATAGCGGCGCTG 322
QY 214 ProLysHisTyrLeuLysTrpLeuAspGlnHisIleHisAlaTyrPheAlaValGlyAla 233
```

```
DB 321 AGCTTCATCACTTCTCTGATTATGACCCACAC-----TTGCCGGA----- 280
QY 234 ProLeuLeuGlySerValGluAlaIleLysSerThrLeuSerGlyValThrPhe----- 251
DB 279 ---ATGAGTGACCGCATCGAAGACGACAGCATACGCTGCGCCCAACTTTCGGTAT 223
QY 252 ---GlyLeuProValSerGluGlyThrAlaArgLeuLeuSerHisSerPheAlaSerSer 270
DB 222 AAGACTTCGCGCTGATATACCAAGACGTGCCCCGCAATTAATACGAATATC----- 175
QY 271 LeuTrpLeuMetProHisSerLysAsnCyseLysGlyAsp-----AsnThrSerTyr 287
DB 174 -----TGCATCGCGCAACTGATCGTTAAACTGCCTG 142
```

Search completed: November 22, 2005, 14:11:21  
Job time : 186 secs

**THIS PAGE BLANK (user)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 10:39:50 ; Search time 4155 Seconds  
(without alignments)  
6148.194 Million cell updates/sec

Title: US-09-651-651-5  
Perfect score: 2965  
Sequence: 1 GCANSKSVTASFTVIAVFL.....DSESIPIKRTAVWELDKSGV 546

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPRO.spool/US09651651/runat\_18112005\_101831\_20261/app.query.fasta\_1.711  
-DB=EST -QPM=faetcap -SUFFX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09651651 @CGN 1.1 4015 @runat\_18112005\_101831\_20261 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOBURY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hlc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1953	65.9	1719	10	CL949010 OaIFSB002
2	1219	41.1	946	7	CR259394 EST705472
3	1122	37.8	903	8	DR929539 EST112107
4	1076	36.3	752	7	CF844688 p8H8029XF
5	1042	35.1	788	7	CR259393 EST705471
6	1007.5	34.0	832	8	DR688001 EST107808
7	965.5	32.6	823	8	DN503413 Y014D01.5

8	952.5	32.1	713	7	CK939714	CK939714 CGF100474
9	919	31.0	686	5	BQ508623	BQ508623 EST616038
10	916.5	30.9	777	8	DR914958	DR914958 EST110649
11	916	30.9	968	7	CK260764	CK260764 EST706842
12	906	30.6	792	2	BG600021	BG600021 EST504916
13	893.5	30.1	704	3	BI422757	BI422757 EST533423
14	877	29.6	596	1	AU237072	AU237072 AU237072
15	841.5	28.4	632	3	BI922233	BI922233 EST542136
16	834	28.1	694	7	CV478732	CV478732 L08B4 F1A
17	832	27.8	676	6	CA090564	CA090564 SCSGAM210
18	823	27.1	565	5	BO584361	BO584361 E011859-0
19	818	27.6	560	6	CD710878	CD710878 VVA026C02
20	810	27.3	728	1	AU306500	AU306500 AU306500
21	808	27.3	683	6	CA221313	CA221313 SCSCF0403
22	795	26.8	628	2	BI267124	BI267124 NF097C071
23	748.5	25.2	647	6	CA156623	CA156623 SCEPR2308
24	734	24.8	534	1	AI974291	AI974291 T110293E
25	730	24.6	543	3	BP954469	BP954469 BP954469
26	715	24.1	478	7	BP886704	BP886704 BP886704
27	689.5	23.3	544	7	CV458559	CV458559 psm01-25m
28	673.5	22.7	617	8	CV526779	CV526779 B13DNF26E
29	601.5	20.3	928	7	CV530907	CV530907 ALV_015B
30	573.5	19.3	797	8	DR440656	DR440656 EST149_35
31	569	18.2	488	2	BI267156	BI267156 NF101G041
32	566	18.2	383	8	CK195018	CK195018 14-B01125
33	550	18.5	502	7	CK115712	CK115712 Y014D01.P
34	541	18.2	394	1	AU290926	AU290926 AU290926
35	523	17.6	518	2	BE321377	BE321377 NF042H091
36	521.5	17.6	625	2	BH944186	BH944186 CB908082
37	489	16.5	792	9	BG367066	BG367066 HYSME1001
38	488	16.5	749	8	DN207859	DN207859 MEST861.B
39	472.5	15.9	417	7	CV189962	CV189962 she01-4ms
40	456	15.4	764	8	DN212885	DN212885 MEST949.B
41	437	14.7	763	8	CV958640	CV958640 pXpxc_63
42	433	14.6	559	1	AW760400	AW760400 g149G03.Y
43	424	14.3	692	7	CO532456	CO532456 3530_1_21
44	413	13.9	682	7	CO521704	CO521704 3530_1_14
45	410	13.8	542	1	AW120335	AW120335 6140B5B09

## ALIGNMENTS

RESULT 1  
LOCUS CL949010 1719 bp DNA linear GSS 21-SEP-2004  
DEFINITION OaIFSB002788 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION CL949010 GI:52361019  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
GSS.  
Oryza sativa (indica cultivar-group)  
Oryza sativa (indica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehmerioideae; Oryzaceae; Oryza.

REFERENCE  
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,  
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
Wong, G. K. S., Deng, X. W. and Wang, J.

TITLE  
An analysis of transcriptional regulation of the rice genome and  
its comparison to Arabidopsis

JOURNAL  
COMMENT  
Unpublished (2004)  
Department of Bioinformatics  
Chinese Academy of Sciences, Beijing 101300, China

CONTACT: Chen Chen  
Tel.: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

FEATURES  
Source location/Qualifiers  
1..1719

/organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:39946"  
 /clone\_lib="Oryza sativa Express Library"  
 /note="Oryza sativa exon trapped genomic sequences"

## ORIGIN

## Alignment Scores:

Pred. No.:	1,216-201	Length:	1719
Score:	1953.00	Matches:	367
Percent Similarity:	74.02%	Conservative:	49
Best Local Similarity:	65.30%	Mismatches:	82
Query Match:	65.87%	Indels:	64
DB:	10	Gaps:	7

US-09-651-651-5 (1-546) x CL949010 (1-1719)

```

Oy      7  SerValThrAlaSerPheThrValIleAlaValPhePheLeuIleCysGlyGlyArgThr 26
         |||||  |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      40  TCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGG 99
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy     27  AlAlaValGluAAspGluThrGluPheHisGlyAAspTyrSerIleuSerGlyIleIleIle 46
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     100  GGAGCGGATGCGCGCGCGGATTC-----GACTACCGGAGCTTCGGGGATATATATAC 153
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy     47  ProGlyPheAlaSerThrGluLeuArgAlaTrpSerIleuAAspCysProGlyThrPro 66
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     154  CCGGGGTGGCGTGCAGCGAGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGG 213
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy     67  LeuAAspPheAAspProLeuAAspLeuValTrpLeuAAspThrThrIleuLeuSerAlaVal 86
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     214  TTGCACTTCAACCCCTCCAGCTCCGCTGCTGCTGCAACCGCCAGCTTTATGCTGTC 273
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy     87  AAspCysTrpPheIleCysMetValLeuAAspProTyrAAspGlnThrAAspHisProGluCys 106
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db     274  AATTGCTGGCTCAAAATGATGCTGCTGCACCTTACCAACGAGATCATCCAGAGTGC 333
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    107  LysSerArgProAAspSerGlyLeuSerAlaIleThrGluLeuAAspProGlyTyrIleThr 126
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    334  AAGTCAGAGCCTCAGTGGTCTTTCTGCAATTACAGAGTTAAGCCCTGATTATTAATCT 393
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    127  GlyProLeuSerThrValTrpLysGluTrpLeuIleTrpCysValGluPheGlyIleGlu 146
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    394  GGTCCCTCTTCTTCAGGTGGAAAGATGGGTCAATGG----- 432
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    147  AlaAAsnAlaIleValAlaValProTyrAAspTrpArgLeuSerProThrIleuGluGlu 166
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    432  ----- 432
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    167  ArgAAspLeuTyrPheHisIleuLeuLeuThrPheGluThrAlaLeuIleuLysLeuArgGly 186
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    433  -----TTAACTTTGAACCTGCAATTGAACTTCGAGGA 465
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    187  GlyProSerIleValPheAlaHisSerMetGlyAAsnAAsnValPheArgTyrPheLeuGlu 206
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    466  GGGCCTTCTTAGTGGTTTGGTCAATCCATGGTAATATGTGTTTGGTCACTTTCTGGAA 525
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    207  TrpLeuAArgLeuGluIleAlaProLysHisTyrLeuLysTrpLeuAAspGlnHisIleHis 226
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    526  TGGTTGAACCTAAGAAATCCCTCCCAAGCACTACATCCGATGGCTTGACGAACATATATACAT 585
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    227  AlaTyrPheAlaValGlyAlaProLeuLeuGlySerAlaGluAlaIleLysSerThrIleu 246
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    586  GCATACCTTGACAGTGGTGCACCTCTTGGATTCTACTAAGCAGATTAAAGCTGCTCTT 645
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    247  SerGlyValThrPheGlyLeuProValSerGlyGlyThrAlaArgLeuLeuSerAAspSer 266
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    646  TCTGGAGCAACATTTGGTCTTCCAGTACGAGGAGGAACACACAGATTGATTTAATGCA 705
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    267  PheAlaSerSerLeuTrpLeuMetProPheSerIleAAspCysIleGlyAAspAAsnThrSer 286
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    706  TTGGTTTCATCTTATATGCTCATGCTCTTCTCAGAAATATTGCAAAAGCTGATTAATATATAC 765
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Oy    287  TrpThrHisPhe-----SerGlyValAlaValIleAlaValLysAAspLysArgValThrHis 303
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    766  TGGAAACATTTCTTTGAGGGAAGGAGGTTGGCCACCAAGACAGCAA----- 813
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    304  CysAAspGluGluGluTyrGlnSerIleSerGlyTyrTrpProThrAAsnIleIleAAsnIle 323
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    814  TGTGATGAATGATGATATATAGCTGATATCTCAGATGATGCCCAAAACCTTGTCAGCATC 873
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    324  GluIleProSerThrSerValThrGluThr-----AlaLeuValAAsn 337
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    874  GAGGTCTTCATCAGTTTCGAGATACGGAACATACCATCATTAAGATATCAACATGAGAAC 933
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    338  MetThr-----SerMetGlyCysGlyLeuProThrIleLeuSerPheThrAlaArgIleu 356
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    934  ATTAACATCCAGATGATGATGATGGAAGCAACATCATTTGTGATTTCTTGATGAGAGGT 993
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    357  AlaAAspGlyThrLeuPheLysAlaIleGluIlePyrAAspProAAspSerPheArgMetLeu 376
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    994  TCAGATGGTACTCTGTTAAACATAAAGATGGACCTCAGAGCATAGAGGCTTATC 1053
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    377  HisGlnLeuLysLeuLeuTyrHisAAspAAspProValPheAAspProLeuThrProThrGlu 396
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1054  CATCAGCTTGAGAAATGACTACCAAGGATATCCAGTTCTTAATCCCTCAGACCTGGAGG 1113
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    397  ArgProProIleLysAAsnValPheCysIleTyrGlyAlaHisAAspLeuThrGluValGly 416
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1114  AAGCCCAATTAAGATATGTTGTCATATATGGCTTGAATCAAGCTAGGTAGGCT 1173
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    417  TyrTyrPheAlaProSerGlyLysProTyrProAAspAAsnTrpIleIleThrAAspIleIle 436
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1174  TATTATTTTGGACCCAGTGAAGAAACATATCCAGATTAACCTGATTAATATGATATTAT 1233
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    437  TyrGluThrGluGlySerLeu-----Val 444
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1234  TATGAATTTGAAGGCTCTTACTATCAAGCAATTAATATGTAACAAATTTATTTCTT 1293
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    445  SerArgSerGlyThrValValAAspGlyAAsnIleGlyProIleThrGlyAAspGluThrVal 464
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1294  TCCAGATTCAGTAAATTCGTTACTGTGAAGAACCCCAACAAATTCAGAGGAGAGGAAACGGTA 1353
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    465  ProTyrHisSerLeuSerTrpCysLysAAsnTrpLeuGlyProLysValAAsnIleThrMet 484
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1354  TCTCAATTCCTCTCATAGTGCAGAAACCTGGCTTGGCCAAAGTGCAAACTAATCTAGG 1413
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    485  AlaProGlnProGluHisAAspGlySerAAspValHisValGluLeuAAsnValAAspHisGlu 504
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1414  GCTCCGACGACGAACATGATGATCTGATTTAACAACAAAGATAGATGCAACCACTAT 1473
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    505  HisGlySerAAspIleIleAlaAAsnMetThrLysAlaProArgValIleTyrIleThrPhe 524
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1474  GTAAGCGACAGGCAATCTCCCAACATGACGAGAGCTCCACATGTGAATGATTAACCTTAC 1533
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    525  TyrGluAAspSerGluSerIleProGlyLysValArgThrAlaValTrpGluLeuAAspLysSer 544
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1534  TTTGAAGATGCTGAAAGCAATTCGGGATGAGAAACAGCGCTCGGAGCTTGATTAAGCA 1593
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Oy    545  GlyTyr 546
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db    1594  AATTCAC 1599
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

## RESULT 2

CK259394  
 LOCUS  
 DEFINITION

EST705472 potato abiotic stress cDNA library Solanum tuberosum CDNA  
 clone POAB029 5' end, mRNA sequence.

ACCESSION  
 CK259394  
 VERSION  
 CK259394.1 GI:39816372

KEYWORDS  
 EST.  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 946)  
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from abiotic stressed potato tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST705471 EST705473

Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potatoc-array@tigr.org  
 Clones can be requested from the University of Arizona Genomics  
 Institute via http://genome.arizona.edu/orders/  
 Seq primer: CAG GAA ACA GCT ATG ACC.

## FEATURES

source  
 1..946  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /culti\_var="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAB029"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /lab\_host="DH10B-Tona"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.: 7,61e-122 Length: 946  
 Score: 1219.00 Matches: 221  
 Percent Similarity: 80.63% Conservative: 33  
 Best Local Similarity: 70.16% Mismatches: 52  
 Query Match: 41.11% Indels: 9  
 DB: 7 Gaps: 2

US-09-651-651-5 (1-546) x CR259394 (1-946)

QY 227 AATATTTGCTGTTGGAGCTCTCTTGTGCTATCGACAGCTCAAGCAACATTA 61  
 DB 2 GCCATTTTGTGTTGGAGCTCTCTTGTGCTATCGACAGCTCAAGCAACATTA 61  
 QY 247 SerGlyValThrPheGlyLeuProValSerGluGlyThrAlaArgLeuLeuSerAsnSer 266  
 DB 62 TCGGGGCTACATTTGGCTTCCTCGTTTCAAGGGGACAGCTCGCGTGAATGTCATATCA 121  
 QY 267 PheAlaSerSerLeuTyrLeuMetProPheSerTyrAsnGlySerGlyAspAsnThrSer 286  
 DB 122 TTTGGATCTTCAATAGATGATTCACATTTTCAAAATATTTACAACTGATGAATGCAAT 181  
 QY 287 TrpThrHisPheSerGlyValAlaAlaValSerAspGlyValTyrHisCysAspGlu 306  
 DB 182 CGAGACATTTCTCTGTGGA---AACGAAATATCCATCATGCTTACCATTTGGCATGAG 238

QY 307 GUGUUTyrGlnSerTyrSerGlyTyrProThrAsnIleIleAsnIleGluIlePro 326  
 DB 239 CATGAACCTTAATCAAGATCTGTGCTGCCAACAAATATATCATTTGAATTCCT 298  
 QY 327 Ser-----ThrsValThrGluThrAlaLeuValAsnMetThr 339  
 DB 299 TCGATTCGAGGGAATGAGTTTATCCCTCAGTTGTGAAGAACTGCCAACCAACTGTCT 358  
 QY 340 SerMetGluCysGlyLeuProThrLeuLeuSerPheThrAlaArgGluLeuAlaAspGly 359  
 DB 359 GGAATGGAATGTGCAATTCCTCCGACACCAATGTCTCTTTCCTGCGTAATCTTGAACGA 418  
 QY 360 ThrLeuPheValAlaIleGluAspTyrAspProAspSerTyrAspMetLeuHisGlnLeu 379  
 DB 419 ACCCTTTTAAAGCCATAAGAAATTAATGATCCTGACAGTGAAGAAAGCTCTTACCTTTTA 478  
 QY 380 LysLeuLeuTyrHisAspAspProValPheAsnProLeuThrProTyrGluArgProPro 399  
 DB 479 AAAAAGCTTACCATATGATGATCGATTTAAATCCTTAACCTTGGGAAAGACCACT 538  
 QY 400 IleLysAsnValPheCysIleTyrGlyAlaHisLeuLeuThrGluValGlyTyrPhe 419  
 DB 539 CTCAGACATATTCGATTTATGCGGTAGATTGCAAGACTGAGGTTGTTACTATTTT 598  
 QY 420 AlaProSerGlyLeuProTyrProAspAsnTyrIleThrAspIleTyrGluThr 439  
 DB 599 GCACCAAGTGGGAAGCTTATCTGTAAATATGATTAACAGATATGATATGAAATC 658  
 QY 440 GUGUIGSerLeuValSerArgSerGlyThrValValAspGlyAsnIleGlyProIleThr 459  
 DB 659 GAAGGATCTCTGATATCCCGCTAGAGAAATCTGTGAAGAAACCTGGTGCACACAGT 718  
 QY 460 GlyAspGluThrValProTyrHisSerLeuSerTyrCysAlaAsnTyrLeuGlyProLys 479  
 DB 719 GGGATGAGACGGGTGCATATCATCTCCCTCTCTGTGCAAGAAATGCGTGGGCCAAAA 778  
 QY 480 ValAsnIleThrMetCysLysProGlnProGluHisAspGlySerAspValHisValGluLeu 499  
 DB 779 GTAAACATTAACAAGACACACAGCTGACAGCATGAGCTGACAGTCAAGATCATCA 838  
 QY 500 AsnValAspHisGluHisGlySerAspIleIleAlaAsnMetThrLysAlaProArgVal 519  
 DB 839 AATATAGAGATCCACATGATGTAAGATATCATTCACCAATATGACAAAGTTACTCATAT 898  
 QY 519 LysTyrIleThrPheTyrGluAspSerGluSerIleProGly 533  
 DB 899 GAAGTACATACCTATATATGAGATTCGAAAGTTTCCAGGA 941

RESULT 3  
 DR929539 903 bp mRNA linear EST 02-AUG-2005  
 LOCUS EST1121078 Aquilegia cDNA library Aquilegia formosa x Aquilegia  
 DEFINITION pubescens cDNA clone COINX27, mRNA sequence.  
 DR929539  
 ACCESSION DR929539  
 VERSION DR929539.1 GI:71698902  
 SOURCE EST.  
 ORGANISM Aquilegia formosa x Aquilegia pubescens  
 Aquilegia formosa x Aquilegia pubescens  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;  
 Ranunculaceae; Aquilegia.  
 1 (bases 1 to 903)  
 Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,  
 Nordborg, M. and Tomkins, J.  
 Generation of ESTs from Aquilegia  
 Unpublished (2005)  
 Other ESTs: EST1121077  
 COMMENT Contact: Scott Hodges  
 Department of Ecology, Evolution and Marine Biology  
 University of California, Santa Barbara  
 Santa Barbara, CA 93106, USA  
 Tel: 805 893 7813  
 Fax: 805 893 4724



Email: hodge@lifesci.ucsb.edu  
Seq primer: M13 Reverse  
Location/Qualifiers  
FEATURES  
source

1. .903  
/organism="Aquilegia formosa x Aquilegia pubescens"  
/mol\_type="mRNA"  
/db\_xref="taxon:338618"  
/clone="COINX27"  
/tissue\_type="mixed shoot and floral apical meristems,  
flower buds, leaves and roots"  
/lab\_host="DH10B T1 (T1 and T5 phage resistance)"  
/clone\_lib="Aquilegia cDNA library"  
/note="Vector: PCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;  
R2, R3, and R4 lines of Aquilegia Formosa X A. pubescens  
were grown from seed in greenhouses at UC Santa Barbara.  
From these plants three sets of tissue were collected: 1)  
Small flower buds (<10 mm) and very young inflorescences  
(71 & 294 by weight respectively), 2) Medium (7-20 mm) and  
large (at or near anthesis) flower buds (65 & 354 by  
weight respectively) and 3) Shoot apical meristems. A  
fourth set of tissue was collected from plants of A.  
formosa. These plants were grown from seed in sand and at  
approximately 1 month root tissue and leaf tissue of  
various developmental stages were collected (84 & 164 by  
weight respectively). Total RNA was extracted from each  
set of tissue and pooled in the following proportions:  
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled  
total RNA, mRNA was extracted and enriched for full-length  
messages and then normalized with proprietary methods by  
Invitrogen."

ORIGIN

Alignment Scores:

Prod. No.: 2,81e-111 Length: 903  
Score: 1122.00 Matches: 203  
Percent Similarity: 80.13% Conservative: 39  
Best Local Similarity: 67.22% Mismatches: 49  
Query Match: 37.84% Indels: 12  
DB: 8 Gaps: 2

US-09-651-651-5 (1-546) x DR929539 (1-903)

QY 218 LeuYvETrPLeuApsGInhiIsIeH;SAIATyPheAlvAlG1vAlaProLeuLeuGly 237  
Db 1 GTCGAATGGCTTATACCAACATACATGCTATTTTGCACTTGAGACCCCTCTTGTGT 60  
QY 238 SerValGluAlaLeuYserThrLeuSerGlyValThrPheGlyLeuProValSerGlu 257  
Db 61 TCCTCCAAACGGTTAAAGAGACATCTCGAGAAACAATTGGCTTCAGTTTCTGAG 120  
QY 258 GlyThrAlaArgLeuLeuSerAenSerPheAlaSerSerLeuTrPLeuMetProPheSer 277  
Db 121 GGAAACGTTGACGTGATTATTAATTCATTGGTTCTTCGTATGATGATGCCATTCTCA 180  
QY 278 LysAenCysLeuGlyAenAenThrSerTrpThrHisPheSerGlyGlyAlaAlaLysLys 297  
Db 181 AGCTATTGTAGACCCGTAATATGTAATCTGAGACATTCTTTAGAGGAAGACGACGGCC 240  
QY 298 AspLysArgValTyHisCysAenGluGluGlyTyrGlnSerLysTySerGlyTrpPro 317  
Db 241 AATCACAATA--TACCACGTGTGATGATGAGTATCGTTTAAATTAATCTGTGATGGCT 297  
QY 318 ThrAenIleLeuAenIleGluIleProSer----- 327  
Db 298 ACTAAATGTGTGAATATCGAAATTCCTCAACTCGAGGTGCAGCTTAAGAACCTTGTGCA 357  
QY 328 ThrSerValThrGluThrAlaLeuValAenMetThrSerMetGluCysGlyLeuProThr 347  
Db 358 ACATT--TTAGATGAAGAAGACAAACAATGTTTCCACGAGGAATGTGACCTCTACT 416  
QY 348 LeuLeuSerPheThrAlaArgGluLeuAlaAenGlyThrLeuPheLysAlaIleGluAen 367  
Db 417 CAGTTATCTCTTCTGTGAGGAAGTGTGATGAGAACATTATTACAGGCTATCAAGAA 476

QY 368 TyrAenProAenSerLysArgMetLeuHISGlnLeuLysLysLeuTyHisAenAenPro 387  
Db 477 TATGATCCAGATACAGAAAACTATTGTATCACTAGAGAAAGTTATATCATGATGATGCA 536  
QY 388 ValPheAenProLeuThrProTrpGluArgProProIleLysAenValPheCysIleTy 407  
Db 537 GTGTTAAATCCTCTTACCCATCGAGATAGACACCTTTGAAGAACATATTTTGCATATAT 596  
QY 408 GlyAlaHisLysLeuThrGluValGlyTyTrpPheAlaProSerGlyLysProTyPro 427  
Db 597 GGAAATGATTCAAAGACTGAGAGTTGGGTATTAATCTTGCTCGAGTGGCAACATATCT 656  
QY 428 AenAenTrpIleIleThrAenIleIleTyGlnThrGluGlnGlySerLeuValSerArgSer 447  
Db 657 GATTAATGATCATCACTAGATGATTTGTTATTAATAATGAGAGCTTATTTTCAAGCT 716  
QY 448 GlyThrValValAenGlyAenAlaGlyProIleThrGlyAenGlyThrValProTyHis 467  
Db 717 GGAAATTTAGTTAATGGGAATCTCGAGCTTCAAGTGGGATGAGACAGTCCCATATAT 776  
QY 468 SerLeuSerTrpCysLysAenTrpLeuGlyProLysValAenIleThrMetAlaProGln 487  
Db 777 TCTCATCATGATGTCAGAACGCTGGCTGGGTCTAAAGTAACATPAACAAGGCTCTCAG 836  
QY 488 ProGluHisAenGlySerAenValHisValGluLeuAenValAenHisGluHisGlySer 507  
Db 837 TCTGAGCATGACGATCTGATGTACAACTGACCTTAACGTAAGCATCATCAAAATGCT 896  
QY 508 AenIle 509  
Db 897 GATATT 902

RESULT 4  
CF844688 752 bp mRNA linear EST 30-OCT-2003  
LOCUS pSHB029Xp13f USDA-IPAFS:Expression of Phytophthora sojae genes  
DEFINITION during infection and propagation\_9HB Phytophthora sojae cDNA clone  
8HB029f13 5, mRNA sequence.

ACCESSION CF844688.1 GI:38060342  
VERSION  
KEYWORDS  
SOURCE Phytophthora sojae  
ORGANISM Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
Phytophthora.  
REFERENCE 1 (bases 1 to 752)  
AUTHORS Tyler,B.  
TITLE Unpublished (2003)  
JOURNAL  
COMMENT  
Tyler lab  
VBI  
1880 Pratt Dr., Blacksburg, VA 24061, USA  
Tel: 540-231-7318  
Email: bmyler@vt.edu  
PCR Primers  
FORWARD: BK reverse primer  
BACKWARD: BK reverse primer  
Plate: 029 row: F column: 13  
Seq primer: BK reverse primer  
High quality sequence stop: 752.

FEATURES  
source location/Qualifiers  
1. .752  
/organism="Phytophthora sojae"  
/mol\_type="mRNA"  
/db\_xref="taxon:67593"  
/clone="SHB029f13"  
/tissue\_type="mycelium"  
/cell\_line="PF6497"  
/dev\_stage="48 hr. post infection stage"  
/lab\_host="Soybean plant"  
/clone\_lib="USDA-IPAFS:Expression of Phytophthora sojae

ORIGIN genes during infection and propagation. sHB"  
/note="Vector: pBK-CMV, Site\_1: EcoRI, Site\_2: XhoI"

Alignment Scores: 2.24e-106 Length: 752  
Pred. No.: 1076.00 Matches: 194  
Score: 92.98% Conservative: 18  
Percent Similarity: 85.09% Mismatches: 13  
Best Local Similarity: 36.29% Indels: 3  
Query Match: 7 Gaps: 1

US-09-651-651-5 (1-546) x CF844688 (1-752)

```

QY 29 GUAAGPGLUthrgluPheHISGLyAspTyrSerLysLeuSerglyLeileileleProGly 48
DB 77 GAAGACGGCCCGCAGCTC-----GACTACTCGAAGCTCCGGCATTATATCCCGCGC 130
QY 49 PheAlaSerThrgluLeuAlaGlaATPSeRileuaspCySProlyrThrProleuasp 68
DB 131 TTGGCATCCACAGCTCCGAGCATGGTCCATCTCCACTGCTCTCCGCTCGAT 190
QY 69 PheasnProleuaspLeuValTTPLeuaspThrThrlsLeuSergAlaValasnCy 88
DB 191 TTCACCCCTCTCGATTGGTCTGCTGACACCAACACTCTTCTGCTGCAATTGC 250
QY 89 TrpPheLysCysMetValleuaspProTyrasnGlnThraspHisProGluCylsSer 108
DB 251 TGGCTTAAGTGATAGTGTGGATCCCTTCAATGACAGATCATCTCGATGCAATGCC 310
QY 109 ArgProaspSerglyLeuSergAlaIleThrgluLeuaspProGlyrThrlleThGlyPro 128
DB 311 CGTCTGATGATGGTCTTCTGCTGATTACAGAACTTGATCCGGGTATATACAGACT 370
QY 129 LeuSerThrValTTPlySGlUTTPLeuLysTrpCysValGluPheGlyIleGluAlaasn 148
DB 371 CTTTCATCGGTTGGAAAGATGATTAGGGTATGATTGCAATTGGCATGAAAGCTAAT 430
QY 149 AlaIleValAlaValProTyrAspTrpArgLeuSerProThrlsLeuGluAlaArgasp 168
DB 431 GCATTAATTCGCTGCTTCATATGATTGAGATGTCCGCCCTCCAAAGCTTGAAGAGCGGAC 490
QY 169 LeuTyrPheHISLysLeuLysLeuThrPheGluThrAlaIleuLysLeuArgGlyPro 188
DB 491 CTTTACTTTCATAGCTAAATAATCACTTGAACCTGCTTACAAACTTCGGGTGGCGCCC 550
QY 189 SerIleValPheAlaHISerMetClyAsnAsnValPheArgTyrPheLeuGluThrLeu 208
DB 551 TCTTTAGTGTGGTGGCAATTCATTTGGTAAATCATGTTCCTTAATTTCTTGGAGTGTG 610
QY 209 ArgLeuGluIleAlaProlYsHISLysLeuLysTrp-LeuaspGlnHISLISLISLTY 228
DB 611 AAGCTAAGATGGATGCCAANACATTAATTCAGTGGGTGGATCAACATATTCGTGCTTA 670
QY 228 rPheAlaValGlyAlaProleuLeuGlySerValGluAlaIleLysSerThrlsLeuSerg 248
DB 671 TTTTGTGCTTGGAGCTCCCTTCTTGGTGCANTGANAACATTTGAAGCAACACTTTCGG 730
QY 248 yValThrPheGlyLeuProVal 255
DB 731 ATTCACATTGGTCTTCTCTATA 752

```

RESULT 5  
CK259393 788 bp mRNA linear EST 03-AUG-2004  
LOCUS CK259393  
DEFINITION EST1705471 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB029 5' end, mRNA sequence.  
ACCESSION CK259393  
VERSION CK259393.1 GI:39816371  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other\_ESTs: EST705470 EST705472 EST705473  
Contact: Roblin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>  
Seq primer: ATT TGG GTG ACA CTA TAG.

FEATURES  
source  
1..788  
Location/Qualifiers  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultiivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB029"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI, Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

Alignment Scores: 1.25e-102 Length: 788  
Pred. No.: 1042.00 Matches: 189  
Score: 82.20% Conservative: 28  
Percent Similarity: 71.59% Mismatches: 38  
Best Local Similarity: 35.14% Indels: 9  
Query Match: 7 Gaps: 2

US-09-651-651-5 (1-546) x CK259393 (1-788)

```

QY 227 AlaTyrPheAlaValGlyAlaProleuLeuGlySerValGluAlaIleLysSerThrl 246
DB 2 GCCTAATTTGCTGTGGAGCTCCCTTCTTGGTGTCTATCGAGACAGTCAAGCAACATTA 61
QY 247 SerGlyVal-ThrPheGlyLeuProValSerglyGlyThrAlaArgLeuSergAasn 266
DB 62 TCCGGGGTCAACATTTTGCTTCTCTTCCAGAGGGAGACAGCTGCTGATGTTCAATAC 121
QY 266 rPheAlaSerSerLeuTTPLeuMetProPheSerLysAsnCyLysGlyaspAntHrse 286
DB 122 ATTGGATCTTCAATATGATGATTACATTTTCAAAATATGTGTAACATGATATGCATA 181
QY 286 rTrpThrlsPheSerglyGlyAlaValLysAspLysArgValTyrHISCyAspG 306

```

```

Db      182  TCGGAGACATTTCTCTGTTGGA---AACGAAATAATCATTCATGCTACCATTTGCATGA 238
Qy      306  uclugluyrglnserlyserglYTrpProthraenlleleasnillegluileer 326
Db      239  GGATGAATTAATCAAGACTCTGCTGGCCCAACAAATATCAATATTCGAATCTCC 298
Qy      326  oser-----ThrservalThrgluThAlaLeuValAsnMetth 339
Db      299  TTCGATTCAGAGGATGAGGTTTATCCCTCAGTTTGAGAAACTGCCCAACCAACTTGTTC 358
Qy      339  rsermetglucyglYleuProthraeusSerPheThraAlaArgluLeuAlaAspGl 359
Db      359  TGGAAATGCAATGTGGATTCCTCCGACACAATTTCTTTTCTGCTCGTGAAGTCTCGACGG 418
Qy      359  YThrLeuPheLyAlaIleGluaAspYrAspProAspSerLyAsrMetLeuHleGlule 379
Db      419  AACCTTTTCAAGCCATTAAGAAATTAATGATCTCGACAGGAAAGACTCTTTACCTTTT 478
Qy      379  ulvylvleuYrHleAspAspProValPheAsnProLeuThrProTgqluArpProP 399
Db      479  AAAAAAGTCTTACCATGATGATCCGATTTTAAATCCTTAACACCTTGGGAAAGACACC 538
Qy      399  oilelyAsnValPheCyalellyrglyAlaHleuLyThrGlulValGlYTrYrPh 419
Db      539  TCTCAAGAACATATCTCGATTTATGCGCTAGATCGAAGACTGAGGTGGTTACTATTT 598
Qy      419  eAlaProserGlyLyProTyProAspAntPrlleleThraPrlleleYrKluTh 439
Db      599  TGCACCAAGGGAGACCTTATCTGATTAATGATTAACAGATGTGATATATGAAT 658
Qy      439  rgluglyserLeuValSerArserSerGlyThrValValaAspGlYAsnAlaGlYProlelle 459
Db      659  CGAAGATCTCTTATTTCCCGGTACAGAAATCTGTTGAAGAAACCTGCTGCACAAAG 718
Qy      459  rglYAspGluThrValProTyrHisSerLeuSerTrpCyAlaAsnTrpleuGlyProly 479
Db      719  TGGGATGAGACCGTGCATACCATTCCTCTCTGCTGCAAGAAATGGCTGGGGCCAAA 778
Qy      479  sValaenlle 482
Db      779  AGTGAACATA 788

RESULT 6
DR688001 832 bp mRNA linear EST 12-JUL-2005
LOCUS     E10708084 Normalized pine embryo library, lib_D Pinus taeda cDNA
DEFINITION
ACCESSION DR688001
VERSION   DR688001
KEYWORDS  EST.
SOURCE    Pinus taeda (loblolly pine)
ORGANISM  Pinus taeda
REFERENCE 1 (bases 1 to 832)
AUTHORS   Buell,C.R., Zheng,L., Cowles,A. and Cairney,J.
TITLE     Spermatophyta; Coniferales; Coniferales; Pinaceae; Pinus.
JOURNAL   Sequencing of ESTs from loblolly pine embryonic libraries
COMMENT   Unpublished (2004)
          Contact: C. Robin Buell
          Plant Genomics Group
          The Institute for Genomic Research
          9712 Medical Center Dr. Rockville, MD 20850, USA
          Email: rbuell@tigr.org
          This clone is available through TIGR. Please contact pine@tigr.org
          for further information
          Seq primer: TAA TAC GAC TCA CTA TAG GG.
          Location/Qualifiers
            1..832
              /organism="Pinus taeda"
              /mol_type="mRNA"
              /cultiVar="7-56 mother tree, open-pollinated tree from,
              Lyons, Georgia, USA"
              /db_xref="taxon:3352"
FEATURES
source

```

```

/clone="PMAB219"
/issue_type="Whole embryos excised from megagametophytes"
/lab_host="DH10B"
/clone_lib="Normalized pine embryo library, lib_D"
/note="Organ: Zygotic Embryos; Vector: pGADT7-RecAB;
Site 1: EcorI; Site 2: ClaI; Library enriched in
transcripts present in suspension culture of loblolly pine
(Pinus taeda L.), 7-56. Suspension culture was established
from fertilized megagametophytes, according to method of
Pullman et al 2003 (Pullman, G., S. Johnson, G. Peter, J.
Cairney, and N. Xu. 2003. Improving loblolly pine somatic
embryo maturation: Comparison of somatic and zygotic
embryo morphology, germination, and gene expression. Plant
Cell Reports 21:747-758. ) This suspension cell culture,
which contains early developing embryos and
undifferentiated embryogenic tissue, was the source of
the RNA used for this suspension culture library."
Subtracted from this was cDNA made from a pool of RNA from
zygotic embryos and female gametophytes from all stages of
development plus somatic embryos from developmental stages
beyond suspension culture (see Clavetta VT, Morillon R,
Pullman GS, Chrispeels M, Cairney J. 2001. An
aquaglyceroporin is abundantly expressed early in the
development of the suspensor and the embryo proper of
loblolly pine (Pinus taeda L.). Plant Physiol. 127:
1556-1567
[http://www.plantphysiol.org/cgi/content/full/127/4/1556])
".

ORIGIN
Alignment Scores:
Pred. No.:      8,03e-99      Length:      832
Score:          1007.50      Matches:      182
Percent Similarity: 84.05%      Conservative: 34
Beet Local Similarity: 70.82%      Mismatches:  39
Query Match:    33.98%      Indels:      2
DB:              8           Gaps:         1

US-09-651-651-5 (1-546) x DR688001 (1-832)
Qy      1  MetGlyAlaAsnSerLySerValThraIAserPheThraValIleAlaValPhePheLeu 20
Db      61  ATGGCAGCATCATCATCACTGTCACAAAGATCTTATTTGTATACCTCTTGCGG 120
Qy      21  lleCyglYlYArghThraIvalaGluaAspGluThrgluPheHleGlYAspYr--Ser 39
Db      121  ATCTCATTTAGAGCCGCAAGAGCCAGTGTGAAGGGAACGATGGTGGAGTACAGTTCC 180
Qy      40  lyLeuSerGlyllelleleProglyPheAlaSerThrglnLeuArghlatrPserile 59
Db      181  AAGCTTTCTGGAGTTATTAATTCGGGATTTGGCTTCAACCACTGAGACATGGGCATG 240
Qy      60  leuAspCyProTyrThrProleuAspPheAsnProleuAspLeuValTrpleuAspTr 79
Db      241  CTCGATTCGCCCTTTTCTCTCTGTATTTCCACCCGCTGGATTCGTTGGCTCGACACC 300
Qy      80  ThrLyLeuLeuSerAlaValaAsnCySTrPheLyCyMetValleuAspProTyraAsn 99
Db      301  AGAAAGTAATTTCTGTGTGAATTTGTGGCTAAAGTGAATGCTTACCCATACAT 360
Qy      100  GlnThraPheIAspProGluCyAluSerArGrProAspSerGlyLeuSerAlaIleThrglu 119
Db      361  CAAAGACATCTCTAATGCAAGTGGCAGCTGACACAGCTGCTCTATTAACAGAG 420
Qy      120  leuAspProGlyYrYlleThrglyProleuSerThraItrlybgluTrpleuYrTrp 139
Db      421  CTCGACCTCGGTTTATTAACAGTCCACCTTCTTGTTGGAAGAAATGGGTCTAGTGG 480
Qy      140  CyValaGlupheGlylleGluaAsnAlaIleValaIValaProTyraAspTrPAsyleu 159
Db      481  TGTGTGAATTTGGAATGAGCAATTCGAATTTGGCTGCTTATTAAGTGAAGGCTTA 540
Qy      160  SerProThrLyLeuGluluaGrAspLeuYrPheHleLyLeuLyLeuThrPheGlu 179

```

```

Db      541 CCAGCTTCGAGCGCGAAGCGGACCTTACTTCCATTAAGCTTAAGTTCCTTTGAA
      180 ThrAlaIeuYsIeuAaVgIyGlyProSerIleValPheAlaHisSerMetGlyAsn 199
      601 ACAGCACTTAAGCTTCGAGTGGTGCATCATTCGTTGTTGCACATTCATTAAGGGAACAT 660
Qy      200 ValPheArgTyPheIeuGluTrpLeuArgIeuGluIleAlaProIySHisTyLeuLys 219
      661 GTCTTCGATATTTTCGAGTGGTGTGAAGTTGAATTCCTCCCAAGCTCTAATATCAG 720
Db      220 TrpIeu-AspGlnHisIleHisAlaTyPheAlaValGlyAlaProIeuLeuGlySerVa 239
      721 TGCGTGGAGTATGACATCATCTATGCTTATATCGCTGTGGAGACACCTTTCTTGGAGCAGT 780
Qy      239 IGUALalleYserTherIeuSerGlyValThrPheGlyLeuProVal 255
      781 AGAGACCATAAAGACAGCTTGTCTGAGCTTACTTTGGCTTCCTATA 829

RESULT 7
DN503413      823 bp      mRNA      linear      EST 10-MAR-2005
LOCUS      Y014D01.spr Populus infected leaf subextracted cDNA library Populus
DEFINITION      tremula cDNA clone Y014D01 5', mRNA sequence.
ACCESSION      DN503413
VERSION      DN503413.1 GI:60713603
KEYWORDS      EST.
SOURCE      Populus tremula
ORGANISM      Populus tremula
REFERENCE      Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
      1 (bases 1 to 823)
      rosidae; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
      Sterky, F., Bhalerao, R.R., Umeberg, P., Segerman, B., Nilsson, P.,
      Brunner, A.M., Charbonnel, L., Campa, L., Lindvall, J.J., Tandré, K.,
      Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,
      Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, U., and Jansson, S.
      A Populus EST resource for plant functional genomics
      Proc. Natl. Acad. Sci. U.S.A. 101 (38), 13951-13956 (2004)
      15353603
      Other ESTs: Y014D01, Y014D01.3pr
      Contact: Bo Segerman
      Umea Plant Science Center, Department of Plant Physiology
      Umea University
      901 87 Umea, Sweden
      Tel: +46 90 786 5279
      Fax: +46 90 786 6676
      Email: bo.segerman@plantphys.umu.se.
      Location/Qualifiers
      1. 823
      /organism="Populus tremula"
      /mol_type="mRNA"
      /db_xref="taxon:113636"
      /clone="Y014D01"
      /issue_type="Virus/Fungal infected leaf"
      /clone_lib="Populus infected leaf subextracted cDNA
      library"

ORIGIN

Alignment Scores:
Pred. No.:      3,07e-94      Length:      823
Score:      965.50      Matches:      188
Percent Similarity:      73.94%      Conservative:      39
Best Local Similarity:      61.24%      Mismatches:      41
Query Match:      32.56%      Indels:      41
Db:      8      Gaps:      3

US-09-651-651-5 (1-546) x DN503413 (1-823)

Qy      216 HisTyLeuLysTrpLeuAspGlnHisIleHisAlaTyPheAlaValGlyAlaProIeu 235
      1 CATTAATACCAAGTGGCTTGAATGACATTCATGCTTATTTTGCAGTTGGAGCTCTCTT 60

```

```

Qy      236 LeuGlySerValGluAlaIleYsSerThrIeuSerGlyValThrPheGlyLeuProVal 255
      61 CTGGGTGCAATTGAGCAGTGAAGCAACATTTTGGGAACACATTTGGTCTTCCAGTT 120
Qy      256 SerGluGlyThrAlaArgIeuSerAsnSerPheAlaSerSerLeuTrpLeuMetPro 275
      121 TCTGAGGGAACAGCTGATTTGATGATTCATTTCTTTGCTCTTCATTATGATGATGCCA 180
Qy      276 PheSerIyAsnCyAlysGlyAspAsnThrSerTrpHisPheSerGlyGlyAlaAla 295
      181 TTTTCAAGTATGATGTAAGCCGATAATCTTACTGTAGGCAATTTTCCGGGGA---TTC 237
Qy      296 LysLysAspLysArgValTyHisCysAspGluGluGlyGlnSerLysTySerGly 315
      238 AGAAGAGGTCATCACACATATCAATCGAGAGACAGAAATTTGTTAAACATATTCGA 297
Qy      316 TrpProThrAsnIleIleAsnIleGluIleProSer-----Thr 338
      298 TGCCCAACAATATATGACACATAGAAATTCCTTATTGCTGGCTTGGTGCCTATCCG 357
Qy      329 SerValThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeu 348
      358 TCAGTTACCGAGTTAGCTCAACCAACTGTCTAGCATGGAAATGCGGCTTCTTACTCAG 417
Qy      349 LeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheLysAlaIleGluAspTy 368
      418 TTATCTTTTCACGCTGTAATAATATCAGATGGGACCTTTTCAAGCAATTTGAAGATTAT 477
Qy      369 AspProAspSerLysArgMetLeuHisGlnLeuLysLeuTyHisAspAspProVal 388
      478 GAATTCAGATGCAAGAGGCTTTTATACCAATTAAAG----- 513
Qy      389 PheAsnProLeuThrProTrpGluArgProProIleLysAsnValPheCysIleTyGly 408
      513 ----- 513
Qy      409 AlaHisLeuLysThrGluValGlyTyTrpPheAlaProSerGlyLysProTyProAsp 428
      514 -----AAAAAACGCTGGTGTACTATTTTGGACCGGAGTGCAGGCTTACCTGAT 564
Qy      429 AsnTrpIleIleThrAspIleIleTyGluThrGluGlySerLeuValSerArgSerGly 448
      565 AATTGATTCACAAACCATGTCATTATAGACTTGAAGGCTCTGTACTCCAGGTCAAGG 624
Qy      449 ThrValValAspGlyAsnAlaGlyProIleThrGlyAspGluThrValProTyHisSer 468
      625 AACATGTCGAAGGATATCTCGAGCTGCAAGTGGGATATACGATACCATCAATTCCT 684
Qy      469 LeuSerTrpCysLysAsnTrpLeuGlyProLysValAsnIleThrMetAlaProGlnPro 488
      685 CTCCTCTTTGTAAAGAAATTGGCTTGGACAAAGTGAACATTAACAGGCTCTCAATGC- 743
Qy      489 GluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySerAsp 508
      744 AGACATATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 802
Qy      509 IleIleAlaAsnMetThrLys 515
      803 ATAGTTCGCCCATGACAGG 823

RESULT 8
CK939714      713 bp      mRNA      linear      EST 19-MAR-2004
LOCUS      CGF1004746.F04 Developing fruit flavedo at 165 DAFB Citrus sinensis
DEFINITION      cDNA clone F1650002.1IF_F04 5', mRNA sequence.
ACCESSION      CK939714
VERSION      CK939714.1 GI:45452970
KEYWORDS      EST.
SOURCE      Citrus sinensis
ORGANISM      Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
      rosidae; eurosids II; Sapindales; Rutaceae; Citrus.

```



Pred. No.: 2.77e-89 Length: 686  
 Score: 919.00 Matches: 160  
 Percent Similarity: 83.94% Conservative: 23  
 Best Local Similarity: 73.39% Mismatches: 35  
 Query Match: 30.99% Indels: 0  
 DB: 5 Gaps: 0

US-09-651-651-5 (1-546) x BG508623 (1-686)

```

Qy 329 SerValThrgIuThraAlaLeuValAsmMetThsSerMetGluCysGlyLeuProThrlau 348
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 60
Db 1 TCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 60

Qy 349 LeuSerPheThraAlaTgGluLeuAlaAspGlyThrlauPheUysAlaIleGluAspTyr 368
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 120
Db 61 TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 120

Qy 369 AspProAspSerIysAspMetLeuHisGlnLeuUysLeuTyrHisAspAspProVal 388
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 180
Db 121 GATCTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 180

Qy 389 PheAsnProLeuThrlProTgGluArgProProIleUysAsnValPheCysIleTyrGly 408
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 240
Db 181 TTAATATCTCTACAACTTGGGAAAGACCACTCTCAAGAAACATATTCGATTATGCGC 240

Qy 409 AlaHisLeuUysThrgIuValGlyTyrThrlPheAlaProSerGlyUysProTyrProAsp 428
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 300
Db 241 GTAGATTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 300

Qy 429 AsnTrpIleIleThraAspIleIleTyrGluThrgIuGlySerLeuValSerArgSerGly 448
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 360
Db 301 AATGGATTATTAACAGATGATGATATGAAATTCAGAGATCTCTGATTCGCCGTCAGGA 360

Qy 449 ThrValValaAspGlyAsnAlaGlyProIleThrgIuAspGluThraValProTyrHisSer 468
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 420
Db 361 AATCTGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 420

Qy 469 LeuSerTrpCysUysAsnTrpLeuGlyProUysValAsnIleThraMetAlaProGlnPro 488
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 480
Db 421 CTCCTGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 480

Qy 489 GluHisAspGlySerAspValHisValGluLeuAsnValaAspHisGluHisGlySerAsp 508
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 540
Db 481 GAGCATGATGCTCAGCTGTACAGTGCATCAAAATATAGAGCATCAACATGATGAGAT 540

Qy 509 IleIleAlaAsnMetThrlPheAlaProArgValUysTyrIleThrlPheTyrGluAspSer 528
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 600
Db 541 ATCATTCCTCAATATGACAAAGTATGCTACATGAAATGATACCTATATGAGGATTC 600

Qy 529 GluSerIleProGlyUysArgThraAlaValTrpGluLeuAspUysSerGlyTyr 546
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 654
Db 601 GAAAGTTTTCAGGAAGACAGACAGCATTTGGGAGCTTGATTAAGCAAAATCAC 654

RESULT 10
DR914958 777 bp mRNA linear EST 02-AUG-2005
LOCUS DR914958 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone CO1672, mRNA sequence.
ACCESSION DR914958.1 GI:71684321
VERSION DR914958.1
KEYWORDS
SOURCE
ORGANISM
EST.
Aquilegia formosa x Aquilegia pubescens
Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
1 (bases 1 to 777)
Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E.,
Nordborg,M. and Tomkins,J.
Generation of ESTs from Aquilegia
Unpublished (2005)
TITLE
JOURNAL
Other ESTs: EST1106496
COMMENT
Contact: Scott Hodges

```

## FEATURES

## source

Department of Ecology, Evolution and Marine Biology  
 University of California, Santa Barbara  
 Santa Barbara, CA 93106, USA  
 Tel: 805 893 7813  
 Fax: 805 893 4724  
 Email: hodges@lifesci.ucsb.edu  
 Seq primer: M13 Reverse

## Location/Qualifiers

## 1..777

/organism="Aquilegia formosa x Aquilegia pubescens"

/mol\_type="mRNA"

/db\_xref="taxon:338618"

/clone="CO11672"

/tissue\_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"

/lab\_host="DH10B T1 (T1 and T5 phage resistance)"

/clone\_lib="Aquilegia cDNA library"

/note="Vector: pCMV SPORT.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa x A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

## ORIGIN

## Alignment Scores:

Pred. No.: 6.34e-89 Length: 777  
 Score: 916.50 Matches: 160  
 Percent Similarity: 85.22% Conservative: 36  
 Best Local Similarity: 69.57% Mismatches: 27  
 Query Match: 30.91% Indels: 7  
 DB: 8 Gaps: 1

US-09-651-651-5 (1-546) x DR914958 (1-777)

```

Qy 324 GluIleProSerThr-----SerValThrgIuThraAlaLeuVal 336
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 61
Db 2 GAAATTCCTCAATATGACAAAGTATGCTACATGAAATGATACCTATATGAGGATTC 61

Qy 337 AsnMetThsSerMetGluCysGlyLeuProThrlLeuLeuSerPheThraAlaTgGluLeu 356
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 121
Db 62 AATCTTTCAGACAGGAAATGTGACCTCCACTCAGTATATCTTCTGCTAGGAAAGTG 121

Qy 357 AlaAspGlyThrlauPheUysAlaIleGluAspTyrAspProAspSerIysAspMetLeu 376
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 181
Db 122 TCAGATGAAACATATATCAGGGCTATCAAGAAATATATATCAATACAGAAACATATTTG 181

Qy 377 HisGlnLeuUysUysLeuTyrHisAspAspProValPheAsnProLeuThrlProTgGlu 396
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 241
Db 182 TATCACTTAAGAAAGTATATATCAATGATGACGTGTTAAATCTCTTACACATGGGAT 241

Qy 397 ArgProProIleUysAsnValPheCysIleTyrGlyAlaHisLeuUysThrgIuValGly 416
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 301
Db 242 AGACCACTTTCAGAAACATATTTTGCATATATGAAATGATTCAAAGACTGAGGTTGG 301

Qy 417 TyrTrpPheAlaProSerGlyUysProTyrProAspAsnTrpIleIleThraAspIle 436
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 361
Db 302 TATATCTTTCCTCGATGCGCAACCACTATTCGATTAACGATCATGATGATATGTT 361

Qy 437 TyrGluThrgIuGlySerLeuValSerArgSerGlyThraValaAspGlyAsnAlaGly 456
    |||||TTCAGTTGTAAGAACTGCCCAACCACTTGTCTGGAATGGAATGGAATTCGCCACACAA 456

```

Db 362 TATGAAATGAAAGACCTTATTTTCAAGCTCTGAAATTTAGTAAATGGAATCTCGA 421  
Qy 457 ProilethrglyAspGluThrValProTyRHisSerLeuSerTrpCysAlaSerIleu 476  
Db 422 GCTTCAAGGGGATGAGACAGTCCATTAATTTCTATCATGTGCAAGACGTGGCTT 481  
Qy 477 GlyProlyValAsnIleThrMetAlaProGlnProGlnHisAspGlySerAspValHis 496  
Db 482 GGGTCTAAAGTAAACCTAACAGAGGCTCCCTCACTGACATGACGATCTGATGTACAA 541  
Qy 497 ValGluLeuAsnValAspHisGlySerAspIleIleAlaAsnMetThrIleValAla 516  
Db 542 GTGGACCTTAAAGCTTAAGCATCTCAACATGCTGATTTGTTCCAAATATGACGAGGAT 601  
Qy 517 ProAsValIleTyRTrpPheTyRgIuAspSerGluSerIleProGlyIleValArgThr 536  
Db 602 CCTCGAGTAAAGTACATCACTACTACGAAGATTTGAAAGTCTTCAGAGAACAGACA 661  
Qy 537 AlaValITrpgIuLeuAspIleSerGlyTyR 546  
Db 662 GCATTTGGAAAGTTGACAAAGCTAGTCAT 691  
RESULT 11  
CK260764 968 bp mRNA linear EST 03-AUG-2004  
LOCUS EST006842 potato abiotic stress cDNA library Solanum tuberosum cDNA  
DEFINITION Clone POAB552 5' end, mRNA sequence.  
ACCESSION CK260764  
VERSION CK260764  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum (potato)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
Buell, C.R., Hart, A., Zismann, V., Karanicheva, S.A. and Baker, B.  
1 (bases 1 to 968)  
Generation of ESTs from abiotic stressed potato tissue  
Unpublished (2003)  
Other ESTs: EST006843  
Contact: Robin Buell  
The Institute for Genomic Research  
9112 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..968  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB552"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_id="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

## ORIGIN

## Alignment Scores:

Pred. No.:	1,02e-88	Length:	968
Score:	916.00	Matches:	160
Percent Similarity:	84.19%	Conservative:	21
Best Local Similarity:	74.42%	Mismatches:	34
Query Match:	30.89%	Indels:	0
DB:	7	Gaps:	0

US-09-651-651-5 (1-546) x CK260764 (1-968)

Qy 329 SerValIThrGluThrAlaLeuValAsnMetThrSerMetGluCysGlyLeuProThrLeu 348  
Db 19 TCAGTTGTAGAAACCTGCCAAACCAACTGTCTGGAAATGAAATGTGATTCGCCACACA 78  
Qy 349 LeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheValAlaIleGluAspTyR 368  
Db 79 TTGTCTTTTTCGTCGTGGAAGTCTCTGTATGGAACCTTTTCAAGCCATAAAGATTAAT 138  
Qy 369 AspProAspSerIleValArgMetLeuHisGlnLeuValIleValIleAspAspProVal 388  
Db 139 GATCCCTGACAGTGAAGAAGCTCTTCAACCTTTTAAAGAGCTTACCATGATATCCGAT 198  
Qy 389 PheAsnProLeuThrProITrpgIuArgPropProIleValAsnValPheCysIleTyRgIy 408  
Db 199 TTAATCCCTTAACACCTTGGAAAGACCACTTCACAGACATATTCGCATTTATGGC 258  
Qy 409 AlaHisLeuIleThrgIuValGlyTyRTrpPheAlaProSerGlyIleAspProTyRProAsp 428  
Db 259 GTAGATTCGAAGACCTGAGGTTGTTATTTTGGACCAAGTGGGAAGCTTATCCTGAT 318  
Qy 429 AsnTrpIleIleThrAspIleIleTyRgIuThrgIuGlySerLeuValSerArgSerGly 448  
Db 319 AATTGATTAATAACAGATGTATATATGAATCGAAGGATCTCTGATATCCCGGTCAAGA 378  
Qy 449 ThrValIleAspGlyAsnAlaGlyProIleThrgIuAspGlyThrValProTyRHisSer 468  
Db 379 AATCTGTGAAGAAACCTGTGTCAACAAGTGGGATGAACGGTGCATACATTTC 438  
Qy 469 LeuSerTrpCysIleAsnTrpLeuGlyProIleValAsnIleThrMetAlaProGlnPro 488  
Db 439 CTCTCCTGTGCAAGAAATTGGCTGGGGCCAAAGTGAACATTAACAAGACACACAGTCA 498  
Qy 489 GluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHisGlySerAsp 508  
Db 499 GAGCATGATGGCTCAGCTCTATACAGATCAAAATATAGACATCAACATGATGAAGAT 558  
Qy 509 IleIleAlaAsnMetThrIleValAlaProArgValIleTyRTrpPheTyRgIuLeuAsp 528  
Db 559 ATCATTTCCAAATATGCAAGATTAAGTACAAATGAACTACATTAACCATATATAGATTTCT 618  
Qy 529 GluSerIleProGlyIleValArgThrAlaValITrpgIuLeuAspIle 543  
Db 619 GAAAGTTTCCAGAAACGAGAACAGACAGATTTGGAGCTTGATAAA 663  
RESULT 12  
BG600021 792 bp mRNA linear EST 07-MAR-2003  
LOCUS EST054916 cSTS solanum tuberosum cDNA clone cSTS27G13 5' sequence,  
DEFINITION mRNA sequence.  
ACCESSION BG600021  
VERSION BG600021  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 792)  
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemiango, A.,  
Bougill, O., Buell, C.R., Romling, C., Tanksley, S. and Baker, B.  
Generations of ESTs from sprouting potato eyes  
Unpublished (2000)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potceto-aray@tigr.org  
This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: M13P-R.

FEATURES  
Source  
Location/Qualifiers

1..792  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cSTS27C13"  
/tissue\_type="sprouting eyes from tubers"  
/dev\_stage="12-14 weeks post harvest"  
/lab\_host="SOLR"  
/note="Vector: Bluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; Various sizes of sprouting eyes (7mm to 15mm) were  
taken from tubers. The tubers were incubated at 26C in the  
dark for 2-3 weeks prior to sprouting. The eyes were  
frozen in liquid nitrogen immediately upon removal from  
tubers."

ORIGIN

Alignment Scores:  
Pred. No.: 9.17e-88 Length: 792  
Score: 906.00 Matches: 160  
Percent Similarity: 83.56% Conservative: 23  
Best Local Similarity: 73.06% Mismatches: 35  
Query Match: 30.56% Indels: 1  
DB: 2 Gaps: 0

US-09-651-651-5 (1-546) x BG600021 (1-792)

QY 329 SerValThrGluThrAlaLeuValAspMetThrSerMetGluGlyLeuProThrLeu 348  
|||||  
1 TCAGTTGTAGAACTGCCCAACCACTTGTCTGAATGGATGTGATTCCTCGACACAA 60  
349 LeuSerPheThrAlaArgGluLeuAlaAspGlyThrLeuPheValAlaIleGluAspTyr 368  
61 TTGTCCTTTCTCTGCTGGAAGCTCTCGACGGAACCTTTTCAAGCCATTAAGAATTAT 120  
369 AspProAspSerLysArgMetLeuHisGlnLeuLysLysLeuTyrHisAspAspProVal 388  
121 GATCTCAAGCTGAAGAACTCTTTCACCTTTAAAAAGCTTACCATGATGATCCGATT 180  
389 PheAsnProLeuThrProTrrGluArgProProIleLysAsnValPheCysIleTyrGly 408  
181 TTAATATCTCTAACACCTTGGAAGACACCTCTCAAGACATATTCGATTTATAGCC 240  
409 AlaHisLeuLysThrGluValGlyTyrTyrPheAlaProSerGlyLysProTyrProAsp 428  
241 GTAGATTGGAAGAGCTGAGGTGGTTACTATTTCACCAAGTGGGAAGCCTTATCTGAT 300  
429 AsnTrpIleIleThrAspIleIleTyrGluThrGluGlySerLeuValSerArgSerGly 448  
301 AATTGATTTATTAACAGATGATATATGAATTCAGAAAGATCTCTGTAATCCCGGTCAAGA 360  
449 ThrValAlaAspGlyAsnAlaGlyProIleThrGluAspGluThrValProTyrHisSer 468  
361 AATCTGGTTGAAGAAACCTGCTGCAACAAAGTGGGATGAGACGTCCTATCATTC 420

QY 469 LeuSerTrpCysLysAsnTrpLeuGlyProLysValAsnIleThrMetAlaProGlnPro 488  
|||||  
421 CTCCTCGTGAAGATGGCTGGGGCCAAAGATGAATATACAAAGACACACAGTCA 480  
489 GluHisAspGlySerAspValHisValGluLeuAsnValAspHisGluHis-GlySerAs 508  
481 GAGCATGATGAGCTGACGCTGTACAGTGATCAAAATATAGACATCAACATGGGTGAGA 540  
508 PileIleAlaAsnMetThrLysAlaProArgValLysTyrIleThrPheTyrGluAsp 528  
541 TATCATTCCTCAATATGACAAAGTTAGCTCAATGAAGTACATTAATATGAGGATTC 600  
QY 528 rGluSerIleProGlyLysArgThrAlaValTrpGluLeuAspLysSerGlyTyr 546  
601 TGAATGTTTCCAGGAACGACAGACAGCTTTGGAGCTTGATTAAGCAATCAC 655

RESULT 13  
BI422757  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI422757 704 bp mRNA linear EST 16-AUG-2001  
EST533423 tomato callus, TMU Lycopersicon esculentum cDNA clone  
cLec71F15 5' end, mRNA sequence.  
BI422757  
BI422757.1 GI:15197439  
EST.  
Lycopersicon esculentum (Solanum lycopersicum)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 704)  
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,  
Holt, I.B., Liang, F., Upton, J., Graven, M.B., Bowman, C.L., Ahn, S.,  
Romling, C.M., Frazer, C.M., Martin, G.B., Tanksley, S.D. and  
Giovannoni, J.  
Generation of ESTs from tomato callus tissue  
Unpublished (1999)  
CONTACT: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES  
Source  
Location/Qualifiers

1..704  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLec71F15"  
/tissue\_type="callus"  
/dev\_stage="25-40 days old"  
/lab\_host="XJ1-Blue MRF"  
/clone\_lib="tomato callus, TMU"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Giovannoni laboratory; cLec - Copydons  
of seedlings 7-10 days post-germination were excised, cut  
at both ends and placed on MS medium with no selection. cut  
Mixed callus was harvested at 25 and 40 days and included  
undifferentiated masses. Tomato Callus EST Library"

ORIGIN

Alignment Scores:  
Pred. No.: 1.77e-86 Length: 704  
Score: 893.50 Matches: 177  
Percent Similarity: 71.85% Conservative: 17  
Best Local Similarity: 65.56% Mismatches: 25  
Query Match: 30.13% Indels: 53  
DB: 3 Gaps: 2

US-09-651-651-5 (1-546) x BI422757 (1-704)

QY 15 IleAlaValPhePheLeuIleCysGlyArgThrAlaValGluAspGluThrGluPhe 34  
49 GTCGTCAATTTTCATTCCTCAGCT-----ACGCGCGTAACTTGGCGGTGAAGTTC 99

QY 35 HsGlyAspTyrSerIleuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeu 54  
Db 100 GCCGGCGATTAATCGAAGCTGTCGGAATTATATCCAGGATCCCGCTACACACG 159  
QY 55 ArgAlaTrpSerIleuAspCysProTyrThrProLeuAspPheAsnProLeuAspLeu 74  
Db 160 AGAGCATGGCATATTAGCTGCGCTTATCTCTCGATTTCATCTCTCGATTATTA 219  
QY 75 ValTrpLeuAspThrThrIleuLeuSerAlaValAsnCysTrpPheIleCysMetVal 94  
Db 220 GTCGTGCTGCAGACCTAATACTTTCTTCGTGAATGCTGTTGAAGTATGANG 279  
QY 95 LeuAspProTyrAsnGlnThrAspHisProGluCysIleSerArgProAspSerGlyLeu 114  
Db 280 TTGGATCTTTACATCAACAGATCAACAGAGTGAAGCTGAGCTGATGATGCGCTT 339  
QY 115 SerAlaIleThrGluLeuAspProGlyTyrIleThrGlyProLeuSerThrValTrpIle 134  
Db 340 TCTGCATATACGAGAGCTTGATCTGTTATATACAGGCTA----- 380  
QY 135 GluTrpLeuIleTrpCysValGluPheGlyIleGluAlaAsnAlaIleValAlaValPro 154  
Db 380 ----- 380  
QY 155 TyrAspTrpArgLeuSerProThrIleuGluGluArgAspLeuTyrPheHisIleu 174  
Db 380 ----- 380  
QY 175 LysLeuThrPheGluThrAlaLeuLysLeuArgGlyGlyProSerIleValPheAlaHis 194  
Db 381 -----ACTTTAAACTGCTCTTAACTTCGAGTGGAGACTTCATAGTTTCCGCTCAT 434  
QY 195 SerMetGlyAsnAsnValPheArgTyrPheLeuGluTrpLeuArgLeuGluIleAlaPro 214  
Db 435 TCATTAGTAAACATGCTTTAGATCTCTTGAAGCTGTAATTAAGATTGCTCCA 494  
QY 215 LysHisTyrLeuLysTrpLeuAspGlnHisIleIleAlaTyrPheAlaValGlyAlaPro 234  
Db 495 AAGAAGTATGCTGATGATTAGACGATCACAATTCATGCTTATTTGCTGTTGAGGCTCCT 554  
QY 235 LeuLeuGlySerValGluAlaIleLysSerThrLeuSerGlyValThrPheGlyLeuPro 254  
Db 555 CTTCTTGCTGCTATTGAGACAGTCMAAGCAATTCGCGGCTCTACATTGGTCTTCC 614  
QY 255 ValSerGluGlyThrAlaArgLeuLeuSerAsnSerPheAlaSerSerLeuTrpLeuMet 274  
Db 615 GTTTCAGAGGAGACAGTTCGCTGATGCTTACATTTGATCTTATATGAGATGTA 674  
QY 275 ProPheSerIleAsnCysLysGlyAspAsn 284  
Db 675 CCAATTTTCAAAATAT-TGTACCACTGATTAAT 703  
RESULT 14  
LOCUS AU237072 596 bp mRNA linear EST 01-APR-2002  
DEFINITION AU237072 RAF1L5 Arabidopsis thaliana cDNA clone RAF1L5-47-A08 5',  
ACCESSION AU237072  
VERSION AU237072  
KEYWORDS mRNA sequence.  
SOURCE AU237072.1 GI:19876241  
ORGANISM EST.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 596)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, F., Kawai, J.,  
Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A.,  
Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
TITLE Large scale analysis of Arabidopsis full-length cDNA  
JOURNAL Unpublished (2002)

COMMENT Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: maseki@rc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
and XhoI was ligated to modified Lambda FLX-1 vector (Carninci et  
al., submitted for publication) digested with BamHI and SalI. This  
clone is in a modified Bluescript vector. Please visit our web  
site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
details.  
FEATURES  
source location/Qualifiers  
1..596  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAF1L5-47-A08"  
/cissue\_type="mixture of silique and flower"  
/lab\_host="DH10B"  
/clone\_1lb="RAF1L5"  
/note="Site\_1: BamHI, Site\_2: SalI"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.62e-85 Length: 596  
Score: 877.00 Matches: 160  
Percent Similarity: 98.77% Conservative: 0  
Best Local Similarity: 98.77% Mismatches: 2  
Query Match: 29.58% Indels: 0  
Gaps: 0  
DB: 1  
US-09-651-651-5 (1-546) x AU237072 (1-596)  
QY 1 MetGlyAlaAsnSerLysSerValThrAlaSerPheThrValIleAlaValPhePheLeu 20  
Db 110 ATGGAGAGCAATTCCGAATCAGTAAACGGTTCCTTACCGCTATGCGCGTTTCTTG 169  
QY 21 IleCysGlyGlyArgThrAlaValGluAspGluThrGluPheHisGlyAspTyrSerLys 40  
Db 170 ATTGGCGGTGGCCGAATCGCGGTGAGATGAGACCGGATTTACGCGCATCTACGAAG 229  
QY 41 LeuSerGlyIleIleIleProGlyPheAlaSerThrGlnLeuArgAlaTrpSerIleLeu 60  
Db 230 CTATCGGATATATCATTCGCGGATTTGGTGGAGCGACGACGACGCGTGGATCTT 289  
QY 61 AspCysProTyrThrProLeuAspPheAsnProLeuAspLeuValTrpLeuAspThrThr 80  
Db 290 GACTGTCATACCTCCGTTGAGATTCATCCGCTCGACCTGATGGCTAGACACCACT 349  
QY 81 LysLeuLeuSerAlaValAsnCysTrpPheIleCysMetValLeuAspProTyrAsnGln 100  
Db 350 AAGCTTCTTTCGCTGCTCAACTGCTGGTTTAAAGTATGATGCTGATCTTATATCA 409  
QY 101 ThrAspHisProGluCysLysSerArgProAspSerGlyLeuSerAlaIleThrGluLeu 120  
Db 410 ACAGACCATCCGAGAGTAACTCAGGCTGACAGTGGCTTTTTCAGCCATCACAGAAATTG 469  
QY 121 AspProGlyTyrIleThrGlyProLeuSerThrValTrpLysGluTrpLeuIleTrpCys 140  
Db 470 GATCCAGTTATACATTAACAGGCTCTTTTCTTACTGCTGGAAGAAGTGGCTMAAGTGGT 529  
QY 141 ValGluPheGlyIleGluAlaAsnAlaIleValAlaValProTyrAspTrpArgLeuSer 160  
Db 530 GTTGAATTTGGTNTGAACCAATGCAATTCGCTGCTTCATACGATTCGAGATGTC 589  
QY 161 ProThr 162  
Db 590 CCAACC 595  
RESULT 15

BI922233 632 bp mRNA linear EST 17-OCT-2001  
 LOCUS EST542136 tomato callus Lycopersicon esculentum cDNA clone  
 DEFINITION CLEC76P18 5' end, mRNA sequence.  
 ACCESSION BI922233  
 VERSION BI922233.1 GI:16218653  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)  
 ORGANISM Lycopersicon esculentum  
 Lycopodium; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 632)  
 Alcala, V., Vredalov, J., White, R., Vision, T., Karamecheva, S. A.,  
 Tsai, J., Uteback, T., Van Aken, S., Roming, C. M., Fraser, C. M.,  
 Martin, G. B., Tanksley, S. D. and Giovannoni, J.  
 Generation of ESTs from tomato callus tissue (2001)  
 Unpublished (2001)  
 COMMENT  
 TITLE  
 JOURNAL  
 CONTACT: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics  
 Institute  
 Seq primer: T3.  
 FEATURES  
 source  
 location/Qualifiers  
 1..632  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA96"  
 /db\_xref="taxon:4081"  
 /clone="CLEC76P18"  
 /issue\_type="callus"  
 /dev\_stage="25-40 days old"  
 /lab\_host="X11-Blue MRP"  
 /clone\_id="Tomato callus"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Boyce Thompson Institute; sequencing: The  
 Institute for Genomic Research, CIBC - Cotyledons of  
 seedlings 7-10 days post-germination were excised, cut at  
 both ends and placed on MS medium with no selection. Mixed  
 callus was harvested at 25 and 40 days and included  
 undifferentiated masses. Tomato Callus EST library"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.17e-81 Length: 632  
 Score: 841.50 Matches: 165  
 Percent Similarity: 72.24% Conservative: 12  
 Best Local Similarity: 67.35% Mismatches: 17  
 Query Match: 28.38% Indels: 52  
 DB: 3 Gaps: 2

US-09-651-651-5 (1-546) x BI922233 (1-632)  
 QY 15 llea1aValPhepHeu1eCyG1yArGThra1aValG1uaSpG1uThrGluPhe 34  
 Db 49 GTGCTCAATTTTCATTCCTTAGCT-----ACCGCGGTAACCTTGCGCGAGATTG 99  
 QY 35 H1aG1yApyrSerLySLeuSerG1y1le1le1leProG1yPhe1aSerThG1uLeu 54  
 Db 100 GCCGCGCATTAACGAGCTGCGGAAATTATCCAGGATTCGCGCTACCAACTG 159  
 QY 55 Arg1a1rPser1le1euaSpCyProLyThrPro1euaSpPheanPro1euaSpLeu 74  
 Db 160 AGAGCAATGCTCAATTAAGACTGCGCTATTCCTCGAATTCATCTTCGATTTA 219  
 QY 75 ValTrPLeuaSpThrThrLySLeu1euaSer1aVala1nCyStrPheLyScyMetVal 94  
 Db 220 GTCTGGCTCGACACTAACTTCTTCTGCTGAATTCGTGTTGAAGTATATGATG 279  
 QY 95 LeuaSpProLyraa1nG1nThraSpH1aProG1uCyS1ySerArgProaSpSerG1yLeu 114

Db 280 TTGGATCTCTTAACAATAAACAGATCAACAGAGTGAAGTCTGACTATAGTGGCTT 339  
 QY 115 Ser1a1le1eThrG1uLeuaSpProG1yTyrl1eThrG1yProLeuSerThrValTrpLys 134  
 Db 340 TCTGCAATTAACGAGAGCTTATCTGCTGTATATTAACAGGCTA----- 380  
 QY 135 GluTrPLeuLyTrpCySValG1uPheG1y1leG1u1a1a1a1a1eVal1a1a1aPro 154  
 Db 380 ----- 380  
 QY 155 TyrAspTrpArgLeuSerProThrLySLeuGluGluArgaSpLeuTyrrPheH1sLysLeu 174  
 Db 380 ----- 380  
 QY 175 LysLeuThrPheG1uThr1a1e1uLySLeuaRgG1yProSer1leValPhe1a1aH1s 194  
 Db 381 -----ACTTTGAAACTGCTTTAAACTTCGCGGAGACCTTCAATGATTTGCTCAT 434  
 QY 195 SerMetG1yAsn1a1nValPheArgTyrrPheLeuGluTrPLeuaRgLeuG1u1le1a1Pro 214  
 Db 435 TCATTAGGAACAATGCTTTAGATCTTCCTGAGGTGTTGAATTAAGATTCCTCA 494  
 QY 215 LysH1aTyrrLeuLystrPLeuaSpG1nH1a1leH1a1aTyrrPhe1a1aValG1y1a1Pro 234  
 Db 495 AAAGAGTATGCTGATGCTTAGACGATCACTTCACTTATTTGCTGTTGAGACTCCT 554  
 QY 235 LeuLeuG1ySerValG1u1a1a1e1uLySserThrLeuSerG1yValThrPheG1yLeuPro 254  
 Db 555 CTTCCTTGTCCTATTGAGACAGTCAGCAACATTAACGGGCTCACTTGTGCTTCTCT 614  
 QY 255 ValSerG1uG1yThr 259  
 Db 615 GTTTCAGAGGGGACA 629

Search completed: November 22, 2005, 13:50:36  
 Job time : 4170 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2005, 01:39:36 ; Search time 7207 Seconds  
(without alignments)  
11033.025 Million cell updates/sec

Title: US-09-651-651-4  
Perfect score: 1641  
Sequence: 1 atggagcgaattcgaaatc.....ctgataaagtggtactaa 1641

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708223 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ha:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1641	100.0	1641	6	AX090311 Sequence
2	911.8	55.6	2609	8	AF493159 Sequence
3	271.8	16.6	3896	6	BD271623 Novel cla
4	271.8	16.6	3896	6	BD271633 Novel cla
5	271.8	16.6	3896	6	AX037587 Sequence
6	271.8	16.6	3896	6	AX037606 Sequence
7	271.8	16.6	119914	8	AC003027 Arabidops
8	176.2	10.7	328	6	AX090328 Sequence
9	118	7.2	93398	8	AP006356 Lotus cor
10	113.4	6.9	118192	8	AP004120 Oryza sat
11	113.4	6.9	147123	8	AP005691 Oryza sat
12	113.4	6.9	158971	2	AP004882 Oryza sat
13	110.4	6.7	709	6	BD271624 Novel cla
14	110.4	6.7	709	6	BD271634 Novel cla
15	110.4	6.7	709	6	AX037588 Sequence
16	110.4	6.7	709	6	AX037607 Sequence
17	109	6.6	356	6	AX090329 Sequence
18	98.6	6.0	3877	8	AF293457 Zea mays
19	68	4.1	110000	8	CR382131_19 Continuation (20 o

20	59	3.6	110000	8	CR382136_07 Continuation (8 of
21	59	3.6	110000	8	CR382136_08 Continuation (9 of
22	57.8	3.5	110000	8	CR382126_06 Continuation (7 of
23	52.6	3.2	733	8	AF428124 Swietenia
24	49.4	3.0	1986	6	BD271616 Novel cla
25	49.4	3.0	1986	6	BD271625 Novel cla
26	49.4	3.0	1986	6	BD271626 Novel cla
27	49.4	3.0	1986	6	AX037577 Sequence
28	49.4	3.0	1986	6	AX037595 Sequence
29	49.4	3.0	1986	6	AX037597 Sequence
30	49.4	3.0	1986	6	AX090382 Sequence
31	49.4	3.0	2575	8	SCYNR008W AX655393 Sequence
32	49.4	3.0	2575	8	SCYNR008W AX655393 Sequence
33	49.4	3.0	23901	6	SCN201952 AX655393 Sequence
34	47.4	2.9	2000	6	BD271617 Novel cla
35	47.4	2.9	2312	6	BD271627 Novel cla
36	47.4	2.9	2312	6	AX037579 Sequence
37	47.4	2.9	2312	6	AX037599 Sequence
38	47.4	2.9	2312	6	SPBC776 AR546016 Sequence
39	47.4	2.9	42391	8	RATCPB01 L07273 Rattus norv
40	47.2	2.9	2106	6	AR546016 Sequence
41	47.2	2.9	5907	10	RATCPB01 L07273 Rattus norv
42	47.2	2.9	261278	2	AC107586 Rattus no
43	43.4	2.6	110000	8	CR380955_05 Continuation (6 of
44	43.4	2.6	110000	8	AE016819_00 Continuation (6 of
45	43.4	2.6	110000	8	AE016819_01 Continuation (2 of

## ALIGNMENTS

RESULT 1	AX090311	1641 bp	DNA	linear	PAT 21-MAR-2001
LOCUS	AX090311	Sequence 4 from Patent WO0116308.			
DEFINITION	AX090311				
ACCESSION	AX090311.1	GI:13444180			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Arabidopsis thaliana (thale cress)				
REFERENCE	Laesner, M. and van Bennekom, A.				
AUTHORS	Laesner, M. and van Bennekom, A.				
TITLE	Plant sterol acyltransferase				
JOURNAL	Patent: WO 0116308-A 4 08-MAR-2001;				
FEATURES	MONSANTO COMPANY (US)				
SOURCE	location/Qualifiers				
	1. 1641				
	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				
ORIGIN					
Query Match	100.0%;	Score 1641;	DB 6;	Length 1641;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1641;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Oy	1	ATGGAGCGAATTGGAATCAGTACGCGCTTCCTACCGTCATCGCGTTTCTTG	60		
Db	1	ATGGAGCGAATTGGAATCAGTACGCGCTTCCTACCGTCATCGCGTTTCTTG	60		
Oy	61	ATTGCGGTCGCGCAACTGCGGTGAGATGAGACCGAGTTTCACTGCGAAG	120		
Db	61	ATTGCGGTCGCGCAACTGCGGTGAGATGAGACCGAGTTTCACTGCGAAG	120		
Oy	121	CTATCGGATTAATCATTCGCGATTGCGTCGACGAGCTACGAGGTGTCATCTT	180		
Db	121	CTATCGGATTAATCATTCGCGATTGCGTCGACGAGCTACGAGGTGTCATCTT	180		
Oy	181	GACTGTCCATACCTCGGTGAGCTTCATTCGCTGACCTGATGAGACCACT	240		
Db	181	GACTGTCCATACCTCGGTGAGCTTCATTCGCTGACCTGATGAGACCACT	240		

Db	181	GACSTCCATACACCTCCGTTGGACCTTCAATCGCGCTGCACCTGTATGGCTACACACACT	240
Qy	241	AAGCTTCTTTCTGCTGTCACACTGCTGGTTTAAAGTATGTCGTAGTCTTTATATCA	300
Db	241	AAGCTTCTTTCTGCTGTCACACTGCTGGTTTAAAGTATGTCGTAGTCTTTATATCA	300
Qy	301	ACAGACCAATCCGAGTGTAAGTCAAGGCGCTGACAGTGCCTTTCAGCCATCAACAATTG	360
Db	301	ACAGACCAATCCGAGTGTAAGTCAAGGCGCTGACAGTGCCTTTCAGCCATCAACAATTG	360
Qy	361	GATCCAGGTTACATTAACAGGTCCTCTTTCTACTGTCGAAAAGAGTGGCTTAAGTGGT	420
Db	361	GATCCAGGTTACATTAACAGGTCCTCTTTCTACTGTCGAAAAGAGTGGCTTAAGTGGT	420
Qy	421	GTTGAGTTGGTATGAAGCAAAATGCAATGTCGCTGTTCCATGAGATGGGATTTGTCA	480
Db	421	GTTGAGTTGGTATGAAGCAAAATGCAATGTCGCTGTTCCATGAGATGGGATTTGTCA	480
Qy	481	CCAACCAAAATGGAGAGCGTGACCTTTACCTTTCACAAGCTCAAGTGAACCTTTGAACCT	540
Db	481	CCAACCAAAATGGAGAGCGTGACCTTTACCTTTCACAAGCTCAAGTGAACCTTTGAACCT	540
Qy	541	GCTTTAAATCCCGGCGGCGCCCTTCTATAGTATTTGCCATTCATGGGTAAATATGTC	600
Db	541	GCTTTAAATCCCGGCGGCGCCCTTCTATAGTATTTGCCATTCATGGGTAAATATGTC	600
Qy	601	TTTCAGATCTTCTTCGAAATGCGCTGAGGCTAGAAATTGCAACAACAATTATTTGAAGTGG	660
Db	601	TTTCAGATCTTCTTCGAAATGCGCTGAGGCTAGAAATTGCAACAACAATTATTTGAAGTGG	660
Qy	661	CTTGATCAGCATATTCAGTCTTATTTCCGTGTGTGGAGCTCTCTTCTTGGTCTGTGAG	720
Db	661	CTTGATCAGCATATTCAGTCTTATTTCCGTGTGTGGAGCTCTCTTCTTGGTCTGTGAG	720
Qy	721	GCAATCAATCTACTCTCTGCTGTAAAGCTTTGGCGCTTCCGTTTCTGAGGGAACGTCT	780
Db	721	GCAATCAATCTACTCTCTGCTGTAAAGCTTTGGCGCTTCCGTTTCTGAGGGAACGTCT	780
Qy	781	CGGTTGTTCCTCAATTCCTTTTGGCGTGTCAATGTGGCTTAATGCCATTTTCAAGAATTGC	840
Db	781	CGGTTGTTCCTCAATTCCTTTTGGCGTGTCAATGTGGCTTAATGCCATTTTCAAGAATTGC	840
Qy	841	AAGGCTGATACACATCCCTGACGGAATTTTCTGGGGGGTGCAGAAAGAAATPAAGGCG	900
Db	841	AAGGCTGATACACATCCCTGACGGAATTTTCTGGGGGGTGCAGAAAGAAATPAAGGCG	900
Qy	901	GTATPACCACTGTGATGAAGAGGAATATCAATCAAAATATTCGTGCTGGCCGACAAATAT	960
Db	901	GTATPACCACTGTGATGAAGAGGAATATCAATCAAAATATTCGTGCTGGCCGACAAATAT	960
Qy	961	ATTAAATTTGAAATTCCTTCCCTAGCGTTACAGAAACAGCTCTAGTCAACATGACAGC	1020
Db	961	ATTAAATTTGAAATTCCTTCCCTAGCGTTACAGAAACAGCTCTAGTCAACATGACAGC	1020
Qy	1021	ATGGAATGTGGCTTCCACACCTTTTGTCTTTCACAGCCCGTGAACCTAGACAGATGGGACT	1080
Db	1021	ATGGAATGTGGCTTCCACACCTTTTGTCTTTCACAGCCCGTGAACCTAGACAGATGGGACT	1080
Qy	1081	CTTTTCAAGCAATGAAAGACTATGACCCGAGTACGAAAGAGATGTACACCACTTAAAG	1140
Db	1081	CTTTTCAAGCAATGAAAGACTATGACCCGAGTACGAAAGAGATGTACACCACTTAAAG	1140
Qy	1141	AAGTTGTATCAGATGACCCCTGTTTTTAATCTCTGACTCTCTTGGGAAAGAACCACTATA	1200
Db	1141	AAGTTGTATCAGATGACCCCTGTTTTTAATCTCTGACTCTCTTGGGAAAGAACCACTATA	1200
Qy	1201	AAAAATGATTTTGTGATATATGTCCTCATCTPAAGACAGAGGTTGGTTATTACTTTGCC	1260
Db	1201	AAAAATGATTTTGTGATATATGTCCTCATCTPAAGACAGAGGTTGGTTATTACTTTGCC	1260
Qy	1261	CCAAATGGCAAACTTATCTGATATATTTGGATCATACGAGATATATTTAGCAAACTGAA	1320
Db	1261	CCAAATGGCAAACTTATCTGATATATTTGGATCATACGAGATATATTTAGCAAACTGAA	1320

QY	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
QY	AF493159	2609 bp mRNA linear PLN 22-SEP-2004	AF493159			Medicago truncatula (barrel medic)	Medicago truncatula (barrel medic)
Db		Medicago truncatula lecithine cholesterol acyltransferase-like protein mRNA, complete cds.	AF493159			Medicago truncatula	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
QY		1 (bases 1 to 2609)				Noitiel, A., Benveniste, P., Banaas, A., Styenne, S. and Bouvier-Nave, P.	
Db		Expression in yeast of a novel phospholipase A1 cDNA from Arabidopsis thaliana				Eur. J. Biochem. 271 (18), 3752-3764 (2004)	
QY		2 (bases 1 to 2609)				Benveniste, P., Bouvier-Nave, P., Schaller, H. and Noitiel, A.	
Db		Direct Submission				Submitted (15-MAR-2002) Plant Molecular Biology Institute, CNRS, 28 rue Goethe, Strasbourg 67083, France	
QY		location/Qualifiers				1. 2609	
Db		/organism="Medicago truncatula"				/mol_type="mRNA"	
QY		/db_xref="taxon:3880"				/clone_11b="Samuel Roberts Noble Foundation Medicago truncatula insect herbivory library (Korth, K., et al., unpublished)"	
Db		/note="Fabaceae; sequence derived from EST clones BE321377 and B1267156"				82. 1980	
QY		/note="enzyme involved in the metabolism of phospholipide and sterols"				/codon_start=1	
Db		/product="lecithine cholesterol acyltransferase-like protein"				/protein_id="AA077002.1"	
QY		/db_xref="GI:25992001"				/translation="MANKPLPISIIAPLAVAGSGSGSELDYSLSGIIIPGSPAS	
Db		TOLRAWSITLCPYSPIDFENFLDVLMDTTLGLSANVCMLEMLDIPNDHPDCKSR				PDSGLSITLDFEYITIGPLSPSWKKEIKTIEGEBANATIAVPIDKRLSPSWLBER	
QY		DVLFHKLKTLFEETAFKLRGSPSLVFGHSLNNVRFLELVKLEIAKHVYIOWDQHI					

HAYFAVAALIGATETIETATLSGFTFGLPVSEGTARLMNPSFASLMMPPSKYCRAS  
 NKYMHPESGKOVQNTYHCDSEBKESFSPMPKUIINELPSRGEAFYSFSEIPE  
 ANLSMEGGLPTQLSFSRAREYADSPFAIEDYDDSRRLIYQEKSTYSLPILNPLI  
 FMDRPIKNVFCITGSNSKIVGYIFASGPRPDNWIITVVYVEESGLVLRGNLVL  
 EGNPDSISGDSITPNSLSMCKNMLKRTVNIITRAPOKSHNIDSDVOIDLVNHHGDDI  
 VPNNTRFPRVXIITYEDESLSLPGKRTVWELDKANHNINIVSSVLMELMLEMRDI  
 HPDASEFVTRAKRGLPDEDCYMDYKARCAMPEYCEYRVFVGVHVGSCRLRYTS  
 ADLLHLHYL"

## ORIGIN

Query Match 55.6%; Score 911.8; DB 8; Length 2609;  
 Best Local Similarity 75.6%; Pred. No. 1.9e-253;  
 Matches 1167; Conservative 0; Mismatches 352; Indels 24; Gaps 2;

108 CGACTACCGAAGCTATGGGTATATCATTCGGGATTTTGGTGGAGCGACTACGAGC 167  
 168 CGATTACTTAACTCCGGCATTTATATCCAGGCTTCGCTTACTACAGCTTCGAGC 227  
 168 GTGGTCATCTTGAAGTCTGCTATACATCCGTTGGACTTCAATCCGCTGAGCTGTATG 227  
 228 ATGGCTATTTTGAAGTCTGCTATACATCCGCTGAGTTCATCTCTCGATTTGGTTTG 287  
 228 GCTAGACACCACTAAGCTTTCTTCTGCTGCTCAACTGCTGTTTAAAGTATGGTCTAGA 287  
 288 GCTGACACCACTAAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 347  
 288 TCCTTATATCAACAGACATCCGAGTGTATGCTACAGGCTGACAGTGTCTTTGAC 347  
 348 CCTTACATCAACAGATCATCCGATTTGTAATCCGCTGCTGCTGCTGCTGCTGCTGCT 407  
 348 CATCAAGATTTGATCCAGGTTATATACAGGCTCTCTTCTATCTGCTGGAAGAGTG 407  
 408 AATTACAGAGCTTATATCAAGGTTATATACAGGCTCTCTGCTGCTGCTGCTGCTGCT 467  
 408 GCTTAAAGTGTGTGTGATTTGATTAAGCAATGCAATGTGCTGCTGCTGCTGCTGCT 467  
 468 GATTAAAGTGTGTGTGATTTGATTAAGCAATGCAATGTGCTGCTGCTGCTGCTGCT 527  
 468 TTGAGATTTGCAACCAATTTGAGAGGCTGACCTTTACTTCAAGCTCAAGT 527  
 528 TTGAGACTGTACCATCATGCTTGAAGAACGAACTTTTCACTTCAATGCTTAAAT 587  
 528 GACCTTTGAACTGCTTAAATCCGAGGCGGCTCTTATAGTATTTGCCATTTCAAT 587  
 588 GACATTTGAGCTGCTTCAAACTTCGAGTGCGCCCTCTTAGTTTGGTCAATTCAT 647  
 588 GGGTATATATGCTTCAAGATCTTCTGAAATGCTGAGGCTAGAAATTGCAACAAACA 647  
 648 GGGTATATATGCTTCTTATTTCTTGAAGTGTGAAGCTAGAAATTTGCAACAAACA 707  
 648 TTATTTGAAGTGTGATTCAGCATATGCTTATTTGCTGTTGAGAGCTCTCTTCT 707  
 708 TTATATCAATGCTGCTGATCAACATATCATGCTTATTTGCTGTTGCTGCCACTTCT 767  
 708 TGGTCTGTTGAGGCAATCAATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767  
 768 GGGTCCAACTGAACCAATGGAACCACTTTCTGATTCACATTTGGTCTTCCAGTATC 827  
 768 TGAAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827  
 828 TGAAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887  
 828 TTCAAAGATTTGCAAGGCTGATTAACATCTCTGAGAGCAATTTTCTGGGGGTGCTGCAA 887  
 888 TTCAAAGATTTGCAAGGCTGATTAACATCTCTGAGAGCAATTTTCTGGGGGTGCTGCAA 887  
 888 TTAAGATTAAGCGCGTATACCACTGATGATGAAGAGATATCAATCAAAATATTTGGCTG 947  
 948 ---AGTCCAAATACATATCACTGTGATGAGAGAGATTAAAGTCAAACTTTTCTGAGATG 947  
 948 GCCGCAAAATATTAATCAATTTGAAATTTCTTCCACTAGCG----- 988  
 1005 GCCAAAGATTAATCAATTTGAAATTTCTTCAACCGGTGATTTGAAGATATCCCTC 1064

989 --TTACAGAAACAGCTCTAGTCAACATGACCAAGATGGAATGTGGCTTCCACCTTTT 1046  
 1065 ATTTTCAGAAATATCCCGAGGCTCACTGTCTCGGAATGGAATGTGACTTACTCAAT 1124  
 1047 GTCTTTACAGCCCGGTGAACATGAGATGAGATCTTTTCAAGCAATAGAAAGTATAGA 1106  
 1125 ATCTTTCTCAGCTGTGAAATATAGCGATGCTCTCTTTTCAAGCAATAGAAAGTATAGA 1184  
 1107 CCCAGATGCAAGAGATGTTTACACCACTTAAAGAGTTGATCATGATGACCTGTTT 1166  
 1185 CCCAGACGCAAGAGACTCTGTACCAATATAGAAATATCATATCTTGGTGTATCCGTACT 1244  
 1167 TAATCTCTGACTCTCTGGGAGAGACCAACCTATAAAAATGATTTGGCATATATGAGC 1226  
 1245 TAATCACTTACCTTGGGATCGGACCAATAAAAAGTGTCTGCAATTTATGAGCAG 1304  
 1227 TCATCTAAAGACAGAGTGTGTTATTAATCTTGGCCCAAGTGGCAAACTTATCTGATTA 1286  
 1305 TAATTCAGAAAGCAAGTTGGTTACTTGTCTTGGCCCTAGCGCAAGCCGTATCTGATTA 1364  
 1287 TTGATCATCAAGGATATCATTTTACGAAATGAAAGTTCCTCGTCAAGTCTGGAAC 1346  
 1365 CTGATCATTACATATGCTGTTTATGAGTACGAAAGATCTCTAGTCACAAGTCAAGGAAA 1424  
 1347 TGTGTTGATGAGGAACCTGGAACCTATTAAGTGGGATGAGACGGTACCTTCAATTCAC 1406  
 1425 TGTGTTGAAAGGAAACCTGTGATCTATTAAGGAGTGAAGGTTCCATCAATCTCCT 1484  
 1407 CTCTTGGTCAAGAAATGGCTCGGACCTTAAGTTAACTAATAGCTCCCAAGCCAGA 1466  
 1485 ATCTTGGTCAAAATTTGCTGGAACCAAGTGAACATTTCAAGAGTCTCTGATCAGA 1544  
 1467 ACAGATGGAAGGACGATGATGGAACCTAATGTTGATCATGACATGAGTGCAGACAT 1526  
 1545 ACAGATGAGCTCAATGATGACATGATTTGAATGGAACATCATATGAGATGATAT 1604  
 1527 CATAGCTAATGACCAAAAGACCAAGGCTTAATGATTAATTTATGAAGCTCGA 1586  
 1605 TGTTCCAACATGACAAAGTGTCTTGAAGTGAATGATTAATTTATGAAGTTCTGA 1664  
 1587 GAGCATTCGGGGAAGAAACCGAGCTGCGGAGCTTGATTA 1629  
 1665 AAGCTTCCAGAAAGGACAGCAGTTTGGGAGCTGATTA 1707

RESULT 3  
 BD271623  
 LOCUS  
 DEFINITION Novel class enzyme in biosynthesis pathway of triacylglycerol production and recombinant DNA molecule encoding the enzyme.  
 ACCESSION BD271623  
 VERSION JP 2002541783-A/8  
 KEYWORDS Arabidopsis thaliana (chale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana (chale cress)  
 REFERENCE  
 AUTHORS Speidel, A., Strahl, U., Lenman, M., Banas, A., Ronne, H., and Stymne, S.  
 TITLE Novel class enzyme in biosynthesis pathway of triacylglycerol production and recombinant DNA molecule encoding the enzyme  
 JOURNAL Patent: JP 2002541783-A 8 10-DEC-2002;  
 COMMENT  
 OS Arabidopsis thaliana (chale cress)  
 PN JP 2002541783-A/8  
 PD 10-DEC-2002  
 PF 28-MAR-2000 JP 2000609586  
 PR 01-APR-1999 EP 99106656.4, 10-JUN-1999 EP 99111321.8 PR  
 07-FEB-2000 US 60/180687  
 PI ANDERS DAHLQVIST, ULF STAHL, MARIT LENMAN, ANTONI BANAS PI  
 , HANS RONNE, STEN STYME

PC C12N15/09,A01H5/00,C12N1/19,C12N5/10,C12N9/10,C12P7/64// PC  
(C12P7/64,C12R1:645), (C12P7/64,C12R1:91), (C12N15/00,C12N5/00 CC  
Novel class enzyme in biosynthesis pathway  
of triacylglycerol  
CC production  
CC and recombinant DNA molecule encoding the enzyme FH Key  
FT source Location/Qualifiers  
FT 1..3896  
cress),  
/organism="Arabidopsis thaliana (thale cress)"  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"

ORIGIN

Query Match 16.6%; Score 271.8; DB 6; Length 3896;  
Best Local Similarity 80.0%; Pred. No. 2.8e-67;  
Matches 375; Conservative 0; Mismatches 2; Indels 92; Gaps 1;

QY 768 TGAGGAACTGCTGCTGTTGTCCTCAATTCCTTTGCGTCATCTTGCGCTTATGCCATT 827  
DB 1787 TCAGGAACTGCTGCTGTTGTCCTCAATTCCTTTGCGTCATCTTGCGCTTATGCCATT 1846

QY 828 TTCAAAGATTGCAAGGCGTATACACATCTCGACGATTTTCTGGGGGTGCTGCAA 887  
DB 1847 TTCAAAGATTGCAAGGCGTATACACATCTCGACGATTTTCTGGGGGTGCTGCAA 1906

QY 888 GAAAGATTAAGGCGGTATACCACTGTGATGAAGAGAAATATCAATAATTTCTGGCTG 947  
DB 1907 GAAAGATTAAGGCGGTATACCACTGTGATGAAGAGAAATATCAATAATTTCTGGCTG 1966

QY 948 GCCGCAAAATTTATTAACATTTGAATTCCTTCCACTAGC----- 987  
DB 1967 GCCGCAAAATTTATTAACATTTGAATTCCTTCCACTAGCCTGTAGACTCTGTATATGCA 2026

QY 988 ----- 987  
DB 2027 ACTGTAACTAACAAGTTTCAACAAGATGTTCACTCTCATATTTCTGTTCTTTGAT 2086

QY 988 -----GTTACAGAAACAGCTCTAGTCAACATGACAGATGGAATGCGCTT 1035  
DB 2087 GTGTATCATATGATTAACAGAAACAGCTCTAGTCAACATGACAGATGGAATGCGCTT 2146

QY 1036 CCCACCCCTTTGCTTTTCAACAGCCCGGTGAAGTGAAGTGGAGCTTTTCAAGCAATA 1095  
DB 2147 CCCACCCCTTTGCTTTTCAACAGCCCGGTGAAGTGAAGTGGAGCTTTTCAAGCAATA 2206

QY 1096 GAAAGCTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT 1144  
DB 2207 GAAAGCTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT 2255

RESULT 4  
BD271633 3896 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Novel class enzyme in biosynthesis pathway of triacylglycerol  
production and recombinant DNA molecule encoding the enzyme.  
ACCESSION BD271633  
VERSION BD271633.1 GI:33081401  
KEYWORDS UP 2002541783-A/18.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;  
1 (bases 1 to 3896)  
Dahlgvist,A., Stahl,U., Lemman,M., Bana,A., Ronne,H. and Syltme,S.  
Novel class enzyme in biosynthesis pathway of triacylglycerol  
production and recombinant DNA molecule encoding the enzyme  
Patent: JP 2002541783-A 18 10-DEC-2002;  
JOURNAL BASF PLANT SCIENCE GMBH

COMMENT OS Arabidopsis thaliana (thale cress)  
PN JP 2002541783-A/18  
PD 10-DEC-2002  
PF 28-MAR-2000 JP 2000609586  
PR 01-APR-1999 EP 99106656.4.10-JUN-1999 EP 99111321.8 PR  
07-FEB-2000 US 60/180687  
PI ANDERS DAHLQVIST, ULF STAHL, MARIT LENMAN, ANTONI BANNAS PI  
, HANS RONNE, STEN STYME  
PC C12N15/09,A01H5/00,C12N1/19,C12N5/10,C12N9/10,C12P7/64// PC  
(C12P7/64,C12R1:645), (C12P7/64,C12R1:91), (C12N15/00,C12N5/00 CC  
Novel class enzyme in biosynthesis pathway  
of triacylglycerol  
CC production  
CC and recombinant DNA molecule encoding the enzyme FH Key  
FT source Location/Qualifiers  
FT 1..3896  
cress),  
/organism="Arabidopsis thaliana (thale cress)"  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"

ORIGIN

Query Match 16.6%; Score 271.8; DB 6; Length 3896;  
Best Local Similarity 80.0%; Pred. No. 2.8e-67;  
Matches 375; Conservative 0; Mismatches 2; Indels 92; Gaps 1;

QY 768 TGAGGAACTGCTGCTGTTGTCCTCAATTCCTTTGCGTCATCTTGCGCTTATGCCATT 827  
DB 1787 TCAGGAACTGCTGCTGTTGTCCTCAATTCCTTTGCGTCATCTTGCGCTTATGCCATT 1846

QY 828 TTCAAAGATTGCAAGGCGTATACACATCTCGACGATTTTCTGGGGGTGCTGCAA 887  
DB 1847 TTCAAAGATTGCAAGGCGTATACACATCTCGACGATTTTCTGGGGGTGCTGCAA 1906

QY 888 GAAAGATTAAGGCGGTATACCACTGTGATGAAGAGAAATATCAATAATTTCTGGCTG 947  
DB 1907 GAAAGATTAAGGCGGTATACCACTGTGATGAAGAGAAATATCAATAATTTCTGGCTG 1966

QY 948 GCCGCAAAATTTATTAACATTTGAATTCCTTCCACTAGC----- 987  
DB 1967 GCCGCAAAATTTATTAACATTTGAATTCCTTCCACTAGCCTGTAGACTCTGTATATGCA 2026

QY 988 ----- 987  
DB 2027 ACTGTAACTAACAAGTTTCAACAAGATGTTCACTCTCATATTTCTGTTCTTTGAT 2086

QY 988 -----GTTACAGAAACAGCTCTAGTCAACATGACAGATGGAATGCGCTT 1035  
DB 2087 GTGTATCATATGATTAACAGAAACAGCTCTAGTCAACATGACAGATGGAATGCGCTT 2146

QY 1036 CCCACCCCTTTGCTTTTCAACAGCCCGGTGAAGTGAAGTGGAGCTTTTCAAGCAATA 1095  
DB 2147 CCCACCCCTTTGCTTTTCAACAGCCCGGTGAAGTGAAGTGGAGCTTTTCAAGCAATA 2206

QY 1096 GAAAGCTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT 1144  
DB 2207 GAAAGCTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT 2255

RESULT 5  
AX037587 3896 bp DNA linear PAT 16-NOV-2000  
LOCUS  
DEFINITION Sequence 11 from Patent W00060055.  
ACCESSION AX037587  
VERSION AX037587.1 GI:11227006  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;



REFERENCE	AUTHORS	TITLE	JOURNAL
1	Banas,A., Stahl,U., Styenne,S., Lemman,M., Rönne,H. and Dahlqvist,A.	A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant dna molecules encoding these enzymes	Patent: WO 0060095-A 11 12-OCT-2000;
			BAAS PLANT SCIENCE GMBH (DE) ; BANAS ANTONI (PL) ; STRAHL ULF (SE) ; STYENNE STEN (SE) ; LEMMAN MARIT (SE) ; RÖNNE HANS (SE) ; DAHLQVIST ANDERS (SE)
FEATURES	source	Location/Qualifiers	
		1..3896	
		/organism="Arabidopsis thaliana"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3702"	
ORIGIN			
Query Match	16.6%;	Score 271.8;	DB 6; Length 3896;
Best Local Similarity	80.0%;	Pred. No. 2.8e-67;	
Matches 375;	Conservative 0;	Mismatches 2;	Indels 92; Gaps 1;
QY	768	TGAGGAACTGCTGCGTTGTTGCCAATCTTTTTCGTCATTTGCGCTTATGCCATT	827
Db	1787	TCAGGAACTGCTGCGTTGTTGCCAATCTTTTTCGTCATTTGCGCTTATGCCATT	1846
QY	828	TTCAAAGATTGCAAGGGTGATTAACAATCTCTGACGCAATTTTTCGGCGGTCTGCATA	887
Db	1847	TTCAAAGATTGCAAGGGTGATTAACAATCTCTGACGCAATTTTTCGGCGGTCTGCATA	1906
QY	888	GAAAGATTAAGGCGGTATTCACCTGCGAAGGAAGGAATATCAATAAATATTCGGCTG	947
Db	1907	GAAAGATTAAGGCGGTATTCACCTGCGAAGGAAGGAATATCAATAAATATTCGGCTG	1966
QY	948	GCCGACAAATTTATTAACATTGGAATCTCTCCACTAGC-----	987
Db	1967	GCCGACAAATTTATTAACATTGGAATCTCTCCACTAGCGGTATGACTCTGTATATGCA	2026
QY	988	-----	987
Db	2027	ACTGTAACTATCAAAAGTTTCAACCAAGATGTTCACTCTCATTTTCGTTCTTGTAT	2086
QY	988	-----GTTACGAAACAGCTCTTACGCAATGACAGATGGAATGCGCCTT	1033
Db	2087	GTTATTCATGATTAACGAAACAGCTCTTACGCAATGACAGATGGAATGCGCCTT	2146
QY	1036	CCCAACCTTTTGTCTTTTACACAGCCCGTGAACCTGACAGATGGAGCTTTTCAAAGCAATA	10959
Db	2147	CCCAACCTTTTGTCTTTTACACAGCCCGTGAACCTGACAGATGGAGCTTTTCAAAGCAATA	2206
QY	1096	GAAGACTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT	1144
Db	2207	GAAGACTATGACCCAGATAGCAAGAGATGTTTACACAGTTAAAGAGT	2255
RESULT 6			
AX037606	3896 bp	DNA	linear
LOCUS	Sequence 30 from Patent WO0060095.		PAT 16-NOV-2000
DEFINITION	AX037606		
ACCESSION	AX037606.1	GI:11227020	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	Arabidopsis thaliana (thale cress)		
	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE			
AUTHORS	Banas,A., Stahl,U., Styenne,S., Lemman,M., Rönne,H. and Dahlqvist,A.		
TITLE	A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant dna molecules encoding these enzymes		
JOURNAL	Patent: WO 0060095-A 30 12-OCT-2000;		
	BAAS PLANT SCIENCE GMBH (DE) ; BANAS ANTONI (PL) ; STRAHL ULF (SE) ;		

FEATURES	SYTAME STEN (SE)	LENNAN MARIT (SE)	RONNE HANS (SE)	DAHLQVIST ANDERS (SE)
source	Location/Qualifiers	1. .3896	/organism="Arabidopsis thaliana"	/mol_type="unassigned DNA"
			/db_xref="taxon:3702"	
ORIGIN				
Query Match	16.6%	Score 271.8;	DB 6;	Length 3896;
Best Local Similarity	80.0%;	Pred. No. 2.86-67/;		
Matches 375;	Conservative 0;	Mismatches 2;	Indels 92;	Gaps 1;
OY	768 TGAGGGAACCTGCTCGCTGTTGTTGCCAATTTCTTTTGGCGTCTCATTTGGCTTATGCCATT	827		
DB	1787 TCAGGGAACCTGCTCGCTGTTGTTGCCAATTTCTTTTGGCGTCTCATTTGGCTTATGCCATT	1846		
OY	828 TTCAAAGAAATTCGAAGGGTGATTAACAATCTGGAGCATTTTTCGAGGGGTGCTGCAAA	887		
DB	1847 TTCAAAGAAATTCGAAGGGTGATTAACAATCTGGAGCATTTTTCGAGGGGTGCTGCAAA	1906		
OY	888 GAAAGATTAAGCGCTATACCACTGTGTGAAGAAGAAATATCAATCAAAATATTTCTGCTG	947		
DB	1907 GAAAGATTAAGCGCTATACCACTGTGTGAAGAAGAAATATCAATCAAAATATTTCTGCTG	1966		
OY	948 GCGGACAAATTTATTTAATCATTTGAATTTCTTCCACTAGC-----	987		
DB	1967 GCGGACAAATTTATTTAATCATTTGAATTTCTTCCACTAGCGGTAGACTCTGTATATGCA	2026		
OY	988 -----	987		
DB	2027 ACTGTAACTACTAACAAAAGTTTCCACCAAGAAATGTTCACTCTCATATTTCTGTTCTTGTAT	2086		
OY	988 -----GTTACAGAAACAGCTCTAGTCAACATGACACAGCATGATGTGGCTT	1035		
DB	2087 GTGTATTCATCAGTTACAGAAACAGCTCTAGTCAACATGACACAGCATGATGTGGCTT	2146		
OY	1036 CCNACCCCTTTTGTCTTTCAAGCCCGTGAACATGACAGATGGGACTCTTTTCAAAAGCAATA	1095		
DB	2147 CCCACCCCTTTTGTCTTTCAAGCCCGTGAACATGACAGATGGGACTCTTTTCAAAAGCAATA	2206		
OY	1096 GAAACATATGACCCAGATTAACAAGAGATGTTACACAGTTAAAGAAAT	1144		
DB	2207 GAAACATATGACCCAGATTAACAAGAGATGTTACACAGTTAAAGAAAT	2255		
RESULT 7				
AC003027	119914 bp	DNA	linear	PLN 30-OCT-2002
LOCUS				
DEFINITION	Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,			
AC003027				
VERSION	complete sequence.			
KEYWORDS	AC003027.1 GI:4079614			
SOURCE	HTG.			
ORGANISM	Arabidopsis thaliana (thale cress)			
REFERENCE	Arabidopsis thaliana			
AUTHORS	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
JOURNAL	1 (bases 1 to 119914)			
REFERENCE	Federipiel,N.A., Palm,C.U., Conway,A.B., Conn,L., Hansen,N.F., Altarfi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,D., Liu,S., Lucero,S., Schwartz,J., Shinn,P., Toriumi,M., Vybotskaia,V.S., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.			
AUTHORS	Unpublished			
JOURNAL	2 (bases 1 to 119914)			
REFERENCE	Federipiel,N.A., Palm,C.U., Conway,A.B., Kutz,D.B., Conway,A.R., Au,M., Araujo,R., Buehler,E., Dewar,K., Peng,J., Kim,C., Li,Y., Oji,O., Osaboro,B.I., Shinn,P., Sun,H., Toriumi,M., Vybotskaia,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA			

Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA  
3 (bases 1 to 119914)  
REFERENCE  
AUTHORS  
Federapfel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Alcafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Lutos,S., Schwartz,U., Shin,P., Toriumi,M., Wyotekala,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
TITLE  
JOURNAL  
Submitted (30-DEC-1998) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
4 (bases 1 to 119914)  
REFERENCE  
AUTHORS  
Federapfel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Alcafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Lutos,S., Schwartz,U., Shin,P., Toriumi,M., Wyotekala,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
TITLE  
JOURNAL  
Submitted (30-JAN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
On Dec 30, 1998 this sequence version replaced gi:2734094.  
Bases 1-9262 of clone F21M11 overlap with bases 66898-78259 of 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7 (AC002560).  
e-mail for correspondence: arab@sequence.stanford.edu  
Genes with similarity to proteins in the databases are described as 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'.  
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge, <http://genome.stanford.edu/~chris/GENSCAN.html>), Fexa (V.Solovayev & A.Salamov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and NePlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.obs.dtu.dk/NePlantGene.html>).  
FEATURES  
source  
1..119914  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="F21M11"  
1..9262  
/note="Overlap with bases 66898-78259 of 'IGF' BAC clone F21M11."  
1..7000  
/note="Proteins in this region are annotated in the F21M11 entry, AC002411."  
7684..10913  
/gene="F21M11.1"  
join(7684..7866,7965..8027,8122..8160,8250..8381,8458..8728,8875..9893,10003..10256,10344..10410,10501..10551,10685..10781,10885..10913)  
/gene="F21M11.1"  
/note="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AADI0665.1"  
/translation="WAPNLRITKACADKMLIGISRTKTRAPFLKLTNYNNPFIIE DAYVLDAITFDEADASTKKKKEKKKKKKSVATSRGRAPRLVNDDE DMDDEDEPLKRLRSRGRASSSSSSSTNNEDLQPEBEDDQVTEPLVDRE VRNGERGLAMTVNNASPSRSSRLSMPEEVPVVLPAHPMETKVEASALVLYLD EPNIDHKFVIDSTGNCAPMLEMKSNIHVEWDETQIDINDTAMVSPSASIGES SEHKVAASVLAISTSGEAKICLSPATGETTVLHPSMEDLRARAEKCLSKYKI VHPEFVGLFMKDMQSCYIDLAKNSTQLTETETVCDMSKAGDSGAIGPLVYVP ECTISGQGWKAIWNKQITAGEENVEIPWNEINKVSRPYMPSFVPDAPYFS LSSPDSQSCSTICIEDCLASBMSCNCAIGVNGPARYTLDDLKEBFLARISEARD

RKQVLRFCESCEPLBRACKVEILBPCKGLKRGAIKECWFCKCCTKRCGNRYVORGHN KLOVEFFNGKMGJRTLEKLPKGAFTCEYJITPELYORSBEDKTPVLIDAH MGSERLEGDCLDGMFVGNIRPLDNLITIEIPVQETPODHYHAFETT RDIAAMEBLADYCIDPRNDLSLKPFDCLCGSRFCRNKAKSTMTWOLANKA"  
13033..14514  
/gene="F21M11.2"  
join(13033..13488,13717..13926,14139..14514)  
/gene="F21M11.2"  
join(13061..13488,13717..13926,14139..14516)  
/gene="F21M11.2"  
/note="Similar to acid phosphatase; Location of ESTs 110C2T7, gb|742036, and 110C2XP, gb|A1100245"  
/codon\_start=1  
/protein\_id="AADI0666.1"  
/db\_xref="GI:4204285"  
/translation="WDRTRPSTLTASLVGVVNSAGDMNLTNOLRGICSSONGYIS KGITDLKGYESRINRVENHNRKPDVYVPECSYH KDYNTSQYKDDVARTDEVI LHFSSMCSKSCDQMDWLPIDDTLLSTPYHKKGFGEGELGNTSPEDWIOKKR APEAVHMKRLYHDIRGQIKIFLISRKEYLRSAVDNLIQAGYGSNMLRGLDEQ OKEVYQYKSEKRKWLMSGYRWGVMDQWSSPAGCPLPRTFPLPNSIYYVA"  
/complement(14004..16362)  
/gene="F21M11.3"  
complement(join(14004..14312,14565..14641,14787..15831,15906..16006,16327..16362))  
/gene="F21M11.3"  
complement(join(14567..14641,14787..15831,15906..16006,16327..16362))  
/gene="Unknown protein; Location of ESTs 40C3TT, gb|AA728590 and40C3T7, gb|704573"  
/codon\_start=1  
/protein\_id="AADI0667.1"  
/db\_xref="GI:4204286"  
/translation="WLDQISGNHTRHKLVSRTAPRPHLNDRVHIVETIPKASVPIRE ICDAERKCSSTSRKRVTPDSVYKYEHYVSESYLSERKEVSESRGSLKSK TDDITERYASNSGSIYPENHRTYKGRSDDIDEDEPDKQSDIDEBEYISDQFSK DLSLPLEVLTQDIOIGKTEIDSKLRSEVETVDSGRYDQGLNPEVNTQKRSK SKGRTKQKQSKQKENSNIADQEEGRDSSPQTDIDITLSPKPKRIEPRKLRNOE LAVAYLSMTWISTGESSEGCNSAMTYLTPEKLSCTSCYSKPLINHDPRVLCALT, EDIOFSATSTPRKSPSPKSPETPTIIGVGVGWGNRKAIDCGSASSPKGIPNTSKY REDKSVNHSPPFARLEKALINDK"  
complement(16835..20238)  
/gene="F21M11.4"  
complement(join(16835..17185,17274..17392,17491..17656,17795..17885,17982..18079,18175..18361,18504..18604,18705..19049,19134..19349,19439..19791,19862..19967,20188..20238))  
/gene="F21M11.4"  
/note="Hypothetical protein"  
/codon\_start=1  
/protein\_id="AADI0669.1"  
/db\_xref="GI:4204288"  
/translation="MGSSPPAPARAAQVPMILIFLEITCYHYVTNRRLKNDVLSA NLNTPKAVTRDLRPMESVISIYKSLNAAYSVHLPQOITPRDQYKADALNSPFGH GSESDSEWTDKDVSKSGGTDSSRDSPLETSSESDPRPDQWTKQSDHLTLVE PESYDANHTPDSBYTEQAAKNVADISASQPSAAKARI CGSPFIOESSNPETQDP TLRLKSLSDSDPTDYKQNHQMVFSFVWVLYFVLCRSNNLRGVIFCVQQLPK SHTQDPSKRKRDIPASDAMENHLKVPKRENNLMKSDIDCONGCSANSDQISEKS KALFOGTSNITIGCFQOSARVSEATGSLHYSRGRPVGDQIPRSNVIVHARSITEMA POYVYEDPTVYNLAELARGMKICTKCSLKAGALGCFVYSCRSYVIVHASEIISRCR WDYDFTLLCPAHSSVFPENKSGHRVSRAPLPKIPAEICSLQTPATFPAKELVLCG SALSRSKLMESLAVRFNATISRYNMPSTVHVYATDEKGACTRLLKVLGLNGKM IINAAWMKASIKASQPVDEEPFEIQTIDOCQDPKTAARLAETNKKLFEGLFYFF GPYKGYKEDONLVKVAAGTILNTEBELAESGSSNNVNDRSSIIVYVINDIPRGCAL GEETITIMORANDMALASQGSLSVGHVTVLESIAQYKLAHPVIG"  
20752..24647  
/gene="F21M11.5"  
join(20752..20994,21363..21497,21596..21740,21825..21994,22102..22178,22542..22758,22920..23064,23343..23442,23599..23693,24021..24072,24227..24298,24441..24647)  
/gene="F21M11.5"  
/note="Hypothetical protein"  
/codon\_start=1



/organism="Lotus corniculatus var. japonicus"  
/mol\_type="genomic DNA"  
/variety="japonicus"  
/db\_xref="taxon:34305"  
/chromosome="6"  
/clone\_lib="LjT library"  
/note="TM0082b, a part of TAC clone:TM0082.-synonym: Lotus japonicus"

## ORIGIN

Query Match 7.2%; Score 118; DB 8; Length 93398;  
Best Local Similarity 79.9%; Pred. No. 1.7e-22;  
Matches 139; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 522 CAAGTTGACCTTTGAACTGCTTTAAATCTCGTGCGCCCTCATATGATTTGCCCA 581  
DB 32553 CAGATTGACATTGAAACCTGCTTAATTAATCTGAGGTGACCCCTCTTTAGTTTGGCCCA 32494  
Oy 582 TTCAATGGGTAAATATGCTTTCAGATTAATCTTTGGAATGCTGAGGCTAGAAATTCACAC 641  
DB 32493 TTCATTGGGTATACAGTTTTCGTTATTTCTTGAGTGCTTGAAGCTAGAGATAGACACC 32434  
Oy 642 AAAAATTAATTTGAGCTGCTGATCAGATATCCATGCTTATTTTCGCTTTGG 695  
DB 32433 AAAACATTAATTTGCAATGCTGATCAGATATCCATGATATTTTGTCTTGG 32380

RESULT 10  
AP004120/c 118192 bp DNA linear PLN 02-JUN-2004  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
BAC clone:OJ1293\_E04.  
AP004120  
ACCESSION  
AP004120.3 GI:47847857

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Euarthridae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (29-AUG-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@ias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Jun 1, 2004 this sequence version replaced gi:34740243.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), GENESH  
(http://www.softberry.com/), Genemark.hmm  
(http://opal.biology.gatech.edu/genemark/), Glimmer  
(http://www.fcr.org/tcd/glimmer/glmr form.html), RICEHMM  
(http://xgp.dna.affrc.go.jp/RICEHMM/), SplicePredictor  
(http://bioinformatics.laastate.edu/cgi-bin/bp.cgi), SIm4  
(http://www.fcr.org/software/glimmer/), BLASTN and BLASTX. The  
genomic sequence was searched against NCBI Nonredundant Protein  
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
sequence database at RGP or DDBJ. Protein homologies of the coding  
regions were searched against NCBI Nonredundant Protein database  
with BLASTP. ESTs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  
BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is

classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with full-length cDNA or  
EST homology (covering almost the entire length of partial  
sequence) is classified as an 'unknown' protein. A gene predicted  
by two or more gene prediction programs is classified as a  
'hypothetical' protein according to IRGSP standard. A gene  
predicted by a single gene prediction program is also classified as  
a probable 'hypothetical' protein and is included as a  
miscellaneous feature of the sequence.  
The orientation of the sequence is from -21M13 to M13rev of the BAC  
clone. This sequence of OJ1293\_E04 clone has an overlap with  
OSUNBA0053L11 (DDBJ: AP005691) clone at 5' end and with P0643P09  
(DDBJ: AP005111) at 3' end. The sequence was generated by combining  
Monsanto and RGP-Japan sequencing data. Detailed information on  
overlap and assembly quality together with annotation of this entry  
is available at  
http://xgp.dna.affrc.go.jp/GenomeSeq.html.

## FEATURES

## SOURCE

1. 118192  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="OJ1293\_E04"  
/gene="OJ1293\_E04.1"  
/join(<1863..1936,2024..>2468)  
/gene="OJ1293\_E04.1"  
/note="start and end point are not identified"  
/join(1863..1936,2024..2468)  
/gene="OJ1293\_E04.1"  
/note="predicted by Genemark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD21651.1"  
/db\_xref="GI:47847858"  
/translation="MELRQNCVQLGDSKVPKRPDPFPSEVEFLKGYREVCCKI  
PCKMLGNSNAAYIVFVAEDSCGLASVWATVGVGRCSTQVCLDSNRNYYE  
GEIYPODGSVILPQERADGMELEBFPVQENNGEVCFSLVKPKAGMWLNGGL  
VIGIEIRPKTS"  
/join(6482..6988,7071..7653,7919..8032,8112..8419,  
8501..8573,8672..9312)  
/gene="OJ1293\_E04.2"  
/join(6482..6988,7071..7653,7919..8032,8112..8419,  
8501..8573,8672..9312)  
/gene="OJ1293\_E04.2"  
/note="transposase-like"  
11875..13972  
/gene="OJ1293\_E04.3"  
/join(11875..12242,12368..12486,13292..13972)  
/gene="OJ1293\_E04.3"  
/note="supported by full-length cDNA(s): AK060823"  
/join(11953..12242,12368..12486,13292..13962)  
/gene="OJ1293\_E04.3"  
/note="supported by full-length cDNA(s): AK098981"  
/join(11964..12242,12368..12486,13292..13787)  
/gene="OJ1293\_E04.3"  
/note="contains EST(s): AU092429(C50563)  
contains full-length cDNA(s): AK060823,AK098981  
phloem-specific lectin"  
/codon\_start=1  
/product="F-box family protein-like"  
/protein\_id="BAD21652.1"  
/db\_xref="GI:47847859"  
/translation="MREEGGICGIARLPRELSAATSPSPDCAAAVSPAPRA  
ADSRVASFPRNLPDLAGEISAPASAKKEFLRISDPEYLSIDLMKMWLRETG  
AKCYMLARSILVIVIGDTPIYRWRIPLTDSRFAGALIVCWLEIRGSHSKMLSPN  
STVAWVWFIAIDFYGDLAPFOASVSLGSGSTKIVCOSYSDSEDEVENYWPMS  
IGPLRRRRRRDRRLVLDGVTVPQKRTDEMBELWGEFINGERGGEVCFISMETK  
GGMWRGLIVGIEIRLRLKSG"  
/join(15841..15901,16106..16272,16375..16439,16593..16662,

## gene

16966..17067,17165..17368,18290..18363,18462..18545,  
19079..19166,19260..19341,19675..19748,19971..20078,  
20180..20244,20814..20927,21141..21240,21337..21390,  
21469..21825)  
/gene="OJ1293\_E04.4-1"  
/join(15841..15901,16106..16272,16375..16439,16593..16662,  
16996..17067,17165..17368,18290..18363,18462..18545,  
19079..19166,19260..19341,19675..19748,19971..20078,  
20180..20244,20814..20927,21141..21240,21337..21390,  
21469..21825)  
/gene="OJ1293\_E04.4-1"  
/note="supported by full-length cDNA(s) : AK100446"  
15858..20860  
/gene="OJ1293\_E04.4-2"  
/join(15858..15901,16996..17067,17165..17368,18290..18363,  
18462..18545,19079..20078,20814..20927,21141..21240,21337..21390,  
21469..21825)  
/gene="OJ1293\_E04.4-2"  
/note="supported by full-length cDNA(s) : AK066398"  
15858..20860  
/gene="OJ1293\_E04.4-2"  
/note="contains full-length cDNA(s) : AK066398  
non-coding transcript  
probably inactive due to including stop codon(s) in CDS"  
/join(16132..16272,16375..16439,16593..16662,16996..17067,  
17165..17368,18290..18363,18462..18545,19079..19166,  
19260..19341,19675..19748,19971..20078,20180..20244,  
20814..20927,21141..21240,21337..21390,21469..21566)  
/gene="OJ1293\_E04.4-1"  
/note="contains EST(s) : D22281 (C10676), AU102120 (C10676)  
contains full-length cDNA(s) : AK100446, AK058270"  
/codon\_start=1  
/product="glutathione reductase"  
/protein\_id="BAD21653.1"  
/db\_xref="GI:47847860"  
/translation="MARKMLKDEVEVAVTDGSDYVDLFTVIGAGSGVGRSRTSASF  
GAKVAICELPFIPISSDQGGGTCVIRGCPKILVYSGSFSEFEADKAFWEIN  
GDIINFMKRLLENKTOEIVRLNGVORLIGNGVTMIEGASLVDAHTEVETKPPGSK  
ORVTAHILATSGRASORVNIKGELATISDALSLEELPKRAVILGGGYLAVERASI  
WKGAGHYDLFPRKELPLRGPDDEMTVAASLBERGRLHGTGLSELSTADITKY  
VTDKSEIILADVLRATGRTPSQKLNLEAAVEVDNIGALIVDYSTRISVPNTAVG  
DVTNRINILPALMEATCTCTKTFGQPTKPRIOVPCVAFSIPLSVASGEQQALE  
BAKSDVLVYTSFNPKNISIKROKRWKLVVDETDVLAASGCPDAPPIIOGMA  
VALKCGATKATPDSVYGIHPSAAEFVFWRTLTRVSPSSKPKTNL"  
complement(27410..29382)  
/gene="OJ1293\_E04.5"  
complement(join(27410..27558,29301..29382))  
/gene="OJ1293\_E04.5"  
/note="start and end point are not identified"  
complement(join(27410..27558,29301..29382))  
/gene="OJ1293\_E04.5"  
/note="predicted by GENSCAN etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD21654.1"  
/db\_xref="GI:47847861"  
/translation="MELSRHRYWPLVYLMEMKRGVVRMRALSTCGTPTPAASGQS  
LCAENRTPTGNSSSHRRGLHRRRLRQTD"  
31777..33484  
/gene="OJ1293\_E04.6"  
31777..33484  
/gene="OJ1293\_E04.6"  
/note="supported by full-length cDNA(s) : AK107210"  
31845..33236  
/gene="OJ1293\_E04.6"  
/note="contains full-length cDNA(s) : AK107210"  
/codon\_start=1  
/product="putative fatty acid elongase 3-ketoacyl-CoA  
synthase 1"  
/protein\_id="BAD21655.1"  
/db\_xref="GI:47847862"  
/translation="MDLPLVTLILLAAHAAVNAQAAARRRATCYLIDVACHKPSD  
DRKVTTELAGAIIERNKGLPEYRFLKLVYNSGIGHTTSPRVNLVAREDCPTLRD  
ALDEMDPFDDAIVAAVAASPRVDVLLVINSGSFSPSLADRVVVRFGLRDDVM

AVNLGMCAGSAGVSDIARVMTLRPTMALVLTSESCEANWYTGDKSNMGLNLF  
ROGGAALITNDPAFRSAKMEELCVARHIGADDDAAVAHREDAGRGVSLSKA  
LPKAAVAFPTNLORIARITLPAEELIARITLREKRAAGAAKINPKTGSD  
HFCILPGTAVITEARSKSLGDSIDVEPARALIRKNGTSSSLTYIISYAEARLN  
AGDRVLTWTFSGFGKCNSSVWVTKDLADGAMEDCIHDYPANLVNPMYKEFGVND  
LPSSQGGWAPFPF"  
/gene="OJ1293\_E04.7"  
35504..35704  
/gene="OJ1293\_E04.7"  
/note="hypothetical ORF

Query Match 6.9%; Score 113.4; DB 8; Length 118192;  
Beet Local Similarity 77.1%; Pred. No. 3.7e-21;  
Matches 138; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

517 AAGCTCAAGTGAACCTTGAACCTGCTTAAACCTCGTGGCGGCTTATAGTATT 576  
64553 AACTCAGATTAACTTTGAAACTGCATTTGAAACTCGAGAGGCGCTTTAGTATT 64494

OY 577 GCCCATCAAGCGGTAATATGCTCTTCGATATCTTTCGAAATGCGTGAAGTAAATT 636  
Db 64493 GCTCATTCATGGTAATATGCTGTTTGGTCTTCTTGAATGCTTGAAGTAAATTC 64434

OY 637 GCACCAAAACATTATTGAAGTGGCTGATGACATATCATGCTTATTGCTGTGG 695  
Db 64433 GCTCCACGATTCATATCCGATGGTTGACGACAACTATATCATGCTTATTCGAGTTGG 64375

RESULT 11  
AP005691/c  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
BAC clone:OSUNBa0053L11.  
AP005691  
AP005691.3 GI:47848524

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehmerioideae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 Sasaki,T., Matsumoto,T. and Katayose,Y.  
Oryza sativa niponbare (GA3) genomic DNA, chromosome 2, BAC  
clone:OSUNBa0053L11  
Published Only in Database (2002)  
2 (bases 1 to 147123)  
Sasaki,T., Matsumoto,T. and Katayose,Y.  
Submitted (05-SEP-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://jgip.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Jun 1, 2004 this sequence version replaced gi:38564216.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mil.edu/GENSCAN.html), RGENESH  
(http://www.softberry.com/), Genemark.hmm  
(http://opal.biology.gatech.edu/genemark/), Glimmer  
(http://www.cigr.org/tcd/glimmer/glmr form.html), RiceHMM  
(http://jgip.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4  
(http://glabin.cse.psu.edu/html/doc/sim4.html), gap2  
(http://www.cigr.org/software/glimmer/), BLASTN and BLASTX. The  
genomic sequence was searched against NCBI Nonredundant Protein  
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
sequence database at RGP or DBP. Protein homologues of the coding  
regions were searched against NCBI Nonredundant Protein database  
with BLASTP. ESTs represent the identified cDNA sequences using  
BLASTN with the corresponding DBP accession no. and RGP clone ID.  
Full-length cDNAs represent the identified cDNA sequences using  
BLASTN with the corresponding DBP accession no.





/product="hypothetical protein"  
/protein\_id="BAD22382.1"  
/db\_xref="GI:47848530"  
/translation="MEVKKPVNSQICVEGVVEILNIGESLVVVKHGIACRDVEVH  
KDVVVEVVKLAK"  
/gene="OSJNB40053J1.1.9"  
/complement(join(21959..22112,22149..22168))  
/gene="OSJNB40053J1.1.9"  
/note="OSJNB40053J1.1.9"  
/name="hypothetical ORF  
predicted by Glimmer"  
this category is not included in IRGSP standard"

Query Match  
Best Local Similarity 77.1%; Score 113.4; DB 8; Length 147123;  
Matches 138; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 517 AAGCTCAAGTGAACCTTGAACCTGCTTAAACCTCGTGGCGCCCTTCTATAGTATT 576  
DB 12651 AACTGACAGTTAACTTTGAACTGATGAACTCGAGGAGGCGCTTCTTAGTGT 126592

QY 577 GCCCATTCATGGGTAATATGCTTCAGATCTTTGGAATGGCTGAGCTAGAAATT 636  
DB 126591 GCTCATTCATGGGTAATATGCTTCAGATCTTTGGAATGGCTGAGAACTAGAAATC 126532

QY 637 GCACCAAAACATTTATTTGAAGTGGCTGATGACATGATCCTTATTTGGCTGTGG 695  
DB 126531 GCTCCAAACATTTATTCATCGATGCTTGAACGAACTATATCATCTTTGCAAGTTGG 126473

RESULT 12  
AP004882/c  
LOCUS  
DEFINITION  
AP004882 158971 bp DNA linear HTG 23-MAR-2002  
Oryza sativa (japonica cultivar-group) chromosome 2 clone P0511E12,  
\*\*\* SEQUENCING IN PROGRESS \*\*\*.  
ACCESSION  
AP004882  
VERSION  
AP004882.1 GI:19698302  
KEYWORDS  
HTG; HTGS PHASE2.  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1 Sasaki, T., Matsunoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC  
clone: P0511E12  
JOURNAL  
2 Published Only in Database (2002)  
AUTHORS  
Sasaki, T., Matsunoto, T. and Yamamoto, K.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (20-MAR-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tssasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the gaps is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
source  
1..158971  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"

FEATURES  
source  
1..158971  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"

ORIGIN  
/clone="P0511E12"

Query Match  
Best Local Similarity 77.1%; Score 113.4; DB 2; Length 158971;  
Matches 138; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 517 AAGCTCAAGTGAACCTTGAACCTGCTTAAACCTCGTGGCGCCCTTCTATAGTATT 576  
DB 58238 AACTGACAGTTAACTTTGAACTGATGAACTCGAGGAGGCGCTTCTTAGTGT 58179

QY 577 GCCCATTCATGGGTAATATGCTTCAGATCTTTGGAATGGCTGAGCTAGAAATT 636  
DB 58178 GCTCATTCATGGGTAATATGCTTCAGATCTTTGGAATGGCTGAGAACTAGAAATC 58119

QY 637 GCACCAAAACATTTATTTGAAGTGGCTGATGACATGATCCTTATTTGGCTGTGG 695  
DB 58118 GCTCCAAACATTTATTCATCGATGCTTGAACGAACTATATCATCTTTGCAAGTTGG 58060

RESULT 13  
BD271624  
LOCUS  
DEFINITION  
BD271624 709 bp DNA linear PAT 17-JUL-2003  
Novel class enzyme in biosynthesis pathway of triacylglycerol  
production and recombinant DNA molecule encoding the enzyme.  
ACCESSION  
BD271624.1 GI:33081392  
VERSION  
JP 2002541783-A/9.  
KEYWORDS  
Lycopersicon esculentum (tomato)  
SOURCE  
Lycopersicon esculentum (tomato)  
ORGANISM  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
1 (bases 1 to 709)  
Dahlqvist, A., Stahl, U., Lennan, M., Banas, A., Ronne, H. and Styenne, S.  
Novel class enzyme in biosynthesis pathway of triacylglycerol  
production and recombinant DNA molecule encoding the enzyme  
Patent: JP 2002541783-A 9 10-DEC-2002;  
JOURNAL  
BASF PLANT SCIENCE GMBH  
OS Lycopersicon esculentum (tomato)  
PN JP 2002541783-A/9  
PD 10-DEC-2002  
PF 28-MAR-2000 JP 2000609586  
PR 01-APR-1999 EP 99106656.4, 10-JUN-1999 EP 99111321.8 PR  
PT 17-FEB-2000 US 60/180687  
PI ANDERS DAHLQVIST, ULF STAHL, MARI LENNAN, ANTONI BANAS PI  
, HANS RONNE, STEN STYENNE  
PC C12N15/09, A01H5/00, C12N1/19, C12N5/10, C12N9/10, C12P7/64// PC  
(C12P7/64, C12R1:645), (C12P7/64, C12R1:91), C12N15/00, C12N5/00 CC  
Novel class enzyme in biosynthesis pathway  
of triacylglycerol  
CC production  
CC and recombinant DNA molecule encoding the enzyme FH Key  
FT source  
1..709  
Location/Qualifiers  
/organism="Lycopersicon esculentum (tomato)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4081"

FEATURES  
source  
1..709  
Location/Qualifiers  
/organism="Lycopersicon esculentum (tomato)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4081"

ORIGIN  
Query Match  
Best Local Similarity 74.5%; Score 110.4; DB 6; Length 709;  
Matches 152; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

QY 1426 CTCGACCTAAGTAACTAATCAATGAGTCCCGACGAGAACGATGAAGACAGCTA 1485  
DB 1 CTGGGGCCAAAGTGAATCAATCAAGACACACAGTCAAGCATGAT-GTTCAATGTGA 59  
QY 1486 CATGTGAACCTAATGTTGATCATGAGATGGGTCAAGCATCATGCTAATACATGACAAA 1545  
|||||

Db 60 CAAGTGCATCTTAATATATAGACATCAACATGGTGAAGATATCATCCCAATATGACAAAG 119

Qy 1546 GCACCAAGGGTTAAGTACATAACCTTTTATGAAGACTGTAGAGCATTCGGGGAAGAGA 1605

Db 120 TTACCTACAAATGAGATCAATAACCTATATGAGGATTCGAAAGTTTCCAGGACAAAGA 179

Qy 1606 ACCGAGTCTGGAGGCTTATATAA 1629

Db 180 ACAGCAGTTTGGAGCTTGATATAA 203

RESULT 14

LOCUS BD271634 709 bp DNA linear PAT 17-JUL-2003

DEFINITION Novel class enzyme in biosynthesis pathway of triacylglycerol production and recombinant DNA molecule encoding the enzyme.

ACCESSION BD271634

VERSION BD271634.1 GI:33081402

KEYWORDS JP 2002541783-A/19.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 709)

AUTHORS Dahlqvist, A., Stahl, U., Lemman, M., Banas, A., Ronne, H. and Styhne, S.

TITLE Novel class enzyme in biosynthesis pathway of triacylglycerol production and recombinant DNA molecule encoding the enzyme

JOURNAL Patent: JP 2002541783-A 19 10-DEC-2002; BASF PLANT SCIENCE GMBH

COMMENT OS Lycopersicon esculentum (tomato)

PN JP 2002541783-A/19

PD 10-DEC-2002

PF 28-MAR-2000 JP 2006009586

PR 01-APR-1999 EP 99106656.4, 10-JUN-1999 EP 99111321.8 PR 07-FEB-2000 US 60/180687

PI ANDERS DAHLQVIST, ULF STRÄHL, MARIT LEMMAN, ANTONI BANAS PI

HANS RONNE, STYHNE

PC C12N15/09, A01H5/00, C12N1/19, C12N5/10, C12N9/10, C12P7/64// PC (C12P7/64, C12R1:645), (C12P7/64, C12R1:91), C12N15/00, C12N5/00 CC

Novel class enzyme in biosynthesis pathway of triacylglycerol

CC and recombinant DNA molecule encoding the enzyme FH Key

Location/Qualifiers

FT source 1..709

FEATURES

source location/Qualifiers

1..709 /organism="Lycopersicon esculentum"

/mol\_type="genomic DNA"

/db\_xref="taxon:4081"

ORIGIN

Query Match 6.7%; Score 110.4; DB 6; Length 709;

Best Local Similarity 74.5%; Pred. No. 2.3e-20;

Matches 152; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

Qy 1426 CTCGACCTAAAGTTAATACATAATGGCTCCCGACCCAGAACAGATGGAAGGACGTA 1485

Db 1 CTGGGGCCAAAAGTGAACATAACAGACACACAGCTGAGCATGAT-GTTCAATGTA 59

Qy 1486 CATGTGAACCTAAATGTTGATCATGAGCATGGGTGAGACATCATAGCTAATGACAAA 1545

Db 60 CAAGTGCATCTTAATATATAGACATCAACATGGTGAAGATATCATCCCAATATGACAAAG 119

Qy 1546 GCACCAAGGGTTAAGTACATAACCTTTTATGAAGACTGTAGAGCATTCGGGGAAGAGA 1605

Db 120 TTACCTACAAATGAGATCAATAACCTATATGAGGATTCGAAAGTTTCCAGGACAAAGA 179

Qy 1606 ACCGAGTCTGGAGGCTTATATAA 1629

Db 180 ACAGCAGTTTGGAGCTTGATATAA 203

RESULT 15

LOCUS AX037588 709 bp RNA linear PAT 16-NOV-2000

DEFINITION Sequence 12 from Patent W00060095.

ACCESSION AX037588

VERSION AX037588.1 GI:11227007

KEYWORDS

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1

AUTHORS Banas, A., Stahl, U., Styhne, S., Lemman, M., Ronne, H. and Dahlqvist, A.

TITLE A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant dna molecules encoding these enzymes

JOURNAL Patent: WO 0060095-A 12 12-OCT-2000; BASF PLANT SCIENCE GMBH (DE) ; BANAS ANTONI (PL) ; STRÄHL ULF (SE) ; STYHNE STEN (SE) ; LEMMAN MARIT (SE) ; RONNE HANS (SE) ; DAHLQVIST ANDERS (SE)

FEATURES

source location/Qualifiers

1..709 /organism="Lycopersicon esculentum"

/mol\_type="unassigned RNA"

/db\_xref="taxon:4081"

ORIGIN

Query Match 6.7%; Score 110.4; DB 6; Length 709;

Best Local Similarity 74.5%; Pred. No. 2.3e-20;

Matches 152; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

Qy 1426 CTCGACCTAAAGTTAATACATAATGGCTCCCGACCCAGAACAGATGGAAGGACGTA 1485

Db 1 CTGGGGCCAAAAGTGAACATAACAGACACACAGCTGAGCATGAT-GTTCAATGTA 59

Qy 1486 CATGTGAACCTAAATGTTGATCATGAGCATGGGTGAGACATCATAGCTAATGACAAA 1545

Db 60 CAAGTGCATCTTAATATATAGACATCAACATGGTGAAGATATCATCCCAATATGACAAAG 119

Qy 1546 GCACCAAGGGTTAAGTACATAACCTTTTATGAAGACTGTAGAGCATTCGGGGAAGAGA 1605

Db 120 TTACCTACAAATGAGATCAATAACCTATATGAGGATTCGAAAGTTTCCAGGACAAAGA 179

Qy 1606 ACCGAGTCTGGAGGCTTATATAA 1629

Db 180 ACAGCAGTTTGGAGCTTGATATAA 203

Search completed: November 9, 2005, 06:57:18

Job time : 7219 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2005, 01:38:48 ; Search time 927 Seconds  
(without alignment) 10479.283 Million cell updates/sec

Title: US-09-651-651-4

Perfect score: 1641

Sequence: 1 atggagagcgaattcgaatc.....cttgataaagtggtatrraa 1641

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*\n2: geneseqn1990s:\*\n3: geneseqn2000s:\*\n4: geneseqn2001s:\*\n5: geneseqn2001bs:\*\n6: geneseqn2002as:\*\n7: geneseqn2002bs:\*\n8: geneseqn2003as:\*\n9: geneseqn2003bs:\*\n10: geneseqn2003cs:\*\n11: geneseqn2003ds:\*\n12: geneseqn2004as:\*\n13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1641	100.0	1641	5	AA601082 Arabidops
2	1622	98.8	1902	12	ADF47823 Arabidops
3	271.8	16.6	3896	3	AA64438 Arabidops
4	271.8	16.6	3896	3	AA64448 Arabidops
5	176.2	10.7	328	5	AA601096 Corn ster
6	110.4	6.7	709	3	AA64449 Lycopersi
7	110.4	6.7	709	3	AA64439 Lycopersi
8	109	6.6	356	5	AA601097 Corn ster
9	52.6	3.2	1872	12	ADF47831 Schizosac
10	49.4	3.0	1701	12	ADF47817 Schizosac
11	49.4	3.0	1986	3	AA64431 Saccharom
12	49.4	3.0	1986	3	AA64441 Saccharom
13	49.4	3.0	1986	3	AA64440 Saccharom
14	49.4	3.0	1986	5	AA601342 Yeast LCA
15	49.4	3.0	1986	12	ADF47816 Yeast acy
16	49.4	3.0	2000	8	ADA71938 Rice gene
17	47.4	2.9	2300	8	ADA71938 Rice gene
18	47.4	2.9	2312	3	AA64442 Schizosac
19	47.4	2.9	2312	3	AA64443 Schizosac
20	40.6	2.5	110000	2	AAV21209_12 Continuation (13 o

## ALIGNMENTS

21	39.4	2.4	2223	12	ADF47830	ADF47830 Aspergill
22	39.4	2.4	7299	13	ADR64383	ADR64383 Aspergill
23	38.8	2.4	486	4	AAF75507	AAF75507 Polygluta
24	38.2	2.3	2047	12	ADF47829	ADF47829 Aspergill
25	37.6	2.3	473	6	ABV94819	ABV94819 Human pan
26	37.6	2.3	1292	10	ADC78228	ADC78228 Human sec
27	37.6	2.3	2329	13	ADG32012	ADG32012 Human DNA
28	37.6	2.3	3329	10	ADR66490	ADR66490 Human pro
29	37.6	2.3	2329	13	ADR65806	ADR65806 Human pro
30	37.6	2.3	3556	12	ADH22401	ADH22401 Human CDN
31	37.6	2.3	3737	6	AA694885	AA694885 Human DNA
32	37.6	2.3	3791	13	ADP23240	ADP23240 PRO polyP
33	37.6	2.3	11477	4	ABL17946	ABL17946 Drosophill
34	37.6	2.3	12600	4	ABL17952	ABL17952 Drosophill
35	37.6	2.3	12600	4	ABL17952	ABL17952 Drosophill
36	37.6	2.3	12600	4	ABL27368	ABL27368 Drosophill
37	36.4	2.2	296	5	ABV18850	ABV18850 Human pro
38	36.4	2.2	110000	6	ABA90521_16	ABA90521_16 Continuation (17 o
39	36	2.2	2433	9	AAI62907	AAI62907 Rice cDNA
40	36	2.2	3579	3	AAA70099	AAA70099 Plasmodiu
41	35.8	2.2	2016	12	ADF47833	ADF47833 Crepie pa
42	35.8	2.2	2976	4	ABA03653	ABA03653 Murine A-
43	35.8	2.2	4590	5	AAH24065	AAH24065 Yeast AOD
44	35.8	2.2	266145	10	ADBE7477	ADBE7477 Fowlpox v
45	35.6	2.2	2356	6	AA62537	AA62537 cDNA sequ

## RESULT 1

AA601082 standard; DNA; 1641 BP.

AA601082; 31-MAY-2001 (first entry)

Arabidopsis thaliana sterol acyltransferase LCAT2 DNA.

lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase; acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil; nutritional supplement; dairy product; food product; salad dressing; de.

Arabidopsis thaliana.

Key Location/Qualifiers

CDS 1..1641 /tag= a /product= "LCAT2"

MO200116308-A2.

08-MAR-2001.

30-AUG-2000; 2000MO-US023863.

30-AUG-1999; 99US-0152493P.

(MONS ) MONSANTO CO.

Lasener M, Van Eenennaam A;

WPI, 2001-169010/17.

P-PSDB; AAU00459.

New isolated nucleic acid encoding plant lecithin:cholesterol acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase-like polypeptides, for modifying the sterol content and oil production of plants.

Claim 5; Page 77; 127pp; English.

CC The present sequence encodes for Arabidopsis thaliana  
CC leclithin:cholesterol acyltransferase-like 2 (LCAT2). Several novel  
CC polynucleotides encoding the plant sterol acyltransferases LCAT  
CC (AA5001081-AA501104, AA501341) and ACAT (acyl CoA:cholesterol  
CC acyltransferase-like; AA501311-AA501319) are described. A yeast LCAT  
CC related open reading frame, LrOI gene sequence (AA501342), and a rat ACAT  
CC (AA501105) cDNA sequence are also described. The polynucleotides encoding  
CC LCAT or ACAT are used to produce LCAT or ACAT polypeptides. They can also  
CC be used in a recombinant construct to transform a host cell (preferably  
CC of a plant) or a plant. The recombinant construct is used to increase or  
CC decrease the sterol content of the host cell or plant. It can be used to  
CC alter oil production of the cell or plant, preferably by increasing it.  
CC The oil of the plant or the plant itself is used as a food product, or as  
CC nutritional or dietary supplements, or in pharmaceutical compositions for  
CC lowering cholesterol. The oil can be used in foods e.g. margarine,  
CC butter, cooking oil, and dressings e.g. salad dressing, mayonnaise,  
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,  
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and  
CC beverages. The alteration in sterol content and/or composition can also  
CC provide a plant with tolerance to stress and insect damage

Sequence 1641 BP; 438 A; 355 C; 378 G; 470 T; 0 U; 0 Other;

Query Match 100.0%; Score 1641; DB 5; Length 1641;

Beef Local Similarity 100.0%; Pred. No. 0;

Matches 1641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCGAATTCGAAATCAGTAACGGCTTCCTCAACCGTCATCGCGTTTCTTG 60  
DB 1 ATGGAGCGAATTCGAAATCAGTAACGGCTTCCTCAACCGTCATCGCGTTTCTTG 60  
QY 61 ATTTCGGGTGCGCGAACTGCGGTGAGAGTGAAGACCGAGTTTCAACGGCACTACGAG 120  
DB 61 ATTTCGGGTGCGCGAACTGCGGTGAGAGTGAAGACCGAGTTTCAACGGCACTACGAG 120  
QY 61 ATTTCGGGTGCGCGAACTGCGGTGAGAGTGAAGACCGAGTTTCAACGGCACTACGAG 120  
DB 61 ATTTCGGGTGCGCGAACTGCGGTGAGAGTGAAGACCGAGTTTCAACGGCACTACGAG 120  
QY 121 CTAATGGGTATATATATTCGGGATTTGGGTGAGCGCACTACGAGCGTGTGATCTT 180  
DB 121 CTAATGGGTATATATATTCGGGATTTGGGTGAGCGCACTACGAGCGTGTGATCTT 180  
QY 121 CTAATGGGTATATATATTCGGGATTTGGGTGAGCGCACTACGAGCGTGTGATCTT 180  
DB 121 CTAATGGGTATATATATTCGGGATTTGGGTGAGCGCACTACGAGCGTGTGATCTT 180  
QY 181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTGACCTGATAGCTAGACCACT 240  
DB 181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTGACCTGATAGCTAGACCACT 240  
QY 181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTGACCTGATAGCTAGACCACT 240  
DB 181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTGACCTGATAGCTAGACCACT 240  
QY 241 AAGCTTCTTTCTGCTGTCAACTGCTGTGTTAAGTATGATGCTGATCTTATATCA 300  
DB 241 AAGCTTCTTTCTGCTGTCAACTGCTGTGTTAAGTATGATGCTGATCTTATATCA 300  
QY 301 ACAGACCATCCGAGGTGAAGTCAAGGCGCTGACAGTGTCTTCAAGCCATCAAGAAATG 360  
DB 301 ACAGACCATCCGAGGTGAAGTCAAGGCGCTGACAGTGTCTTCAAGCCATCAAGAAATG 360  
QY 361 GATTCAGGTATACATTAACAGTCTCTTCTACTGTCTGGAAGAGTGTGCTTAAGTGT 420  
DB 361 GATTCAGGTATACATTAACAGTCTCTTCTACTGTCTGGAAGAGTGTGCTTAAGTGT 420  
QY 421 GTTGGATTGTATGAAGCAATATGCTGCTGTTCAATGATGGAGATGTGCA 480  
DB 421 GTTGGATTGTATGAAGCAATATGCTGCTGTTCAATGATGGAGATGTGCA 480  
QY 481 CCAACCAATTTGGAAGAGCGTGAACCTTCACTTCAAGCTCAAGTGAACCTTGAACCT 540  
DB 481 CCAACCAATTTGGAAGAGCGTGAACCTTCACTTCAAGCTCAAGTGAACCTTGAACCT 540  
QY 541 GCTTTAAATCTCGTGGCGGCCCTTCTATAGTATTTGCCATTCATGAGGTAATATGTC 600  
DB 541 GCTTTAAATCTCGTGGCGGCCCTTCTATAGTATTTGCCATTCATGAGGTAATATGTC 600  
QY 601 TTCAATATCTTTCTGGAATGCTGAGGCTAGAAATTTGCAACCAATATTTGAAGTGG 660  
DB 601 TTCAATATCTTTCTGGAATGCTGAGGCTAGAAATTTGCAACCAATATTTGAAGTGG 660  
QY 661 CTGTATCGCATATCATGCTTATTTGCTGTGAGGCTCTTCTTGTTCTGTGAG 720  
DB 661 CTGTATCGCATATCATGCTTATTTGCTGTGAGGCTCTTCTTGTTCTGTGAG 720

DB 661 CTGTATCGCATATCATGCTTATTTGCTGTGAGGCTCTTCTTGTTCTGTGAG 720  
QY 721 GCAATCAATATCACTCTCTGTGTGTAAAGTTTGGCTCTCTGTTCTGAGGGAATCTGT 780  
DB 721 GCAATCAATATCACTCTCTGTGTGTAAAGTTTGGCTCTCTGTTCTGAGGGAATCTGT 780  
QY 781 CGGTGTGTCCAAATCTTTTGTGCGTGTATGTGTGCTTATGCAATTTTCAAGAAATGTC 840  
DB 781 CGGTGTGTCCAAATCTTTTGTGCGTGTATGTGTGCTTATGCAATTTTCAAGAAATGTC 840  
QY 841 AAGGTGTATACATCTCTGAGACGATTTTCTGGGGGTGCTGCAAGAAATATAGGC 900  
DB 841 AAGGTGTATACATCTCTGAGACGATTTTCTGGGGGTGCTGCAAGAAATATAGGC 900  
QY 901 GTAATCACTGTGATGAAGAGAAATATCAATCAAAATTTCTGTGCTGCGCAAAATAT 960  
DB 901 GTAATCACTGTGATGAAGAGAAATATCAATCAAAATTTCTGTGCTGCGCAAAATAT 960  
QY 961 ATTAACATTTAAATCTTCACTGACGTTACAGAAACAGCTCTTATGCAATGACAGC 1020  
DB 961 ATTAACATTTAAATCTTCACTGACGTTACAGAAACAGCTCTTATGCAATGACAGC 1020  
QY 1021 ATGGAAATGTGCTTCCCAACCTTTTGTCTTCAACGCGGTGAACATGAGGGAATCT 1080  
DB 1021 ATGGAAATGTGCTTCCCAACCTTTTGTCTTCAACGCGGTGAACATGAGGGAATCT 1080  
QY 1081 CTTTCAAGAGCAATGAAGATATGACCCAGATAGCAAGAGATTTACACAGTTAAAG 1140  
DB 1081 CTTTCAAGAGCAATGAAGATATGACCCAGATAGCAAGAGATTTACACAGTTAAAG 1140  
QY 1141 AAGTTGATCATGATGACCTGTTTATATCTCTGACTCTTGGAGAGACCACTATA 1200  
DB 1141 AAGTTGATCATGATGACCTGTTTATATCTCTGACTCTTGGAGAGACCACTATA 1200  
QY 1201 AAAAATGATTTTGTGATATATGTGCTCATATCAAGAGAGGTTGTTTACTTTGCC 1260  
DB 1201 AAAAATGATTTTGTGATATATGTGCTCATATCAAGAGAGGTTGTTTACTTTGCC 1260  
QY 1261 CCAAGTGGCAAACTTATCTGATATATGATATCAACGATATCATATTAGAAACTGAA 1320  
DB 1261 CCAAGTGGCAAACTTATCTGATATATGATATCAACGATATCATATTAGAAACTGAA 1320  
QY 1321 GGTTCCTCTGTGTCAAGTCTGGAACGTGTGTTAGTGGAAACGCTGACCTTATACTGGG 1380  
DB 1321 GGTTCCTCTGTGTCAAGTCTGGAACGTGTGTTAGTGGAAACGCTGACCTTATACTGGG 1380  
QY 1381 GATGAGAGGTAACTTATCTGATATATGATATCAACGATATCATATTAGAAACTGAA 1440  
DB 1381 GATGAGAGGTAACTTATCTGATATATGATATCAACGATATCATATTAGAAACTGAA 1440  
QY 1441 AACATTAACATGTGCTCCCAAGCAGAAACAGATGGAAGCGATATGTGGAATTAAT 1500  
DB 1441 AACATTAACATGTGCTCCCAAGCAGAAACAGATGGAAGCGATATGTGGAATTAAT 1500  
QY 1501 GTTGTATGATGATGCTGCTGAGACATCTATGATCAATGACAAAGCAACCAAGGTTAAG 1560  
DB 1501 GTTGTATGATGATGCTGCTGAGACATCTATGATCAATGACAAAGCAACCAAGGTTAAG 1560  
QY 1561 TACATTAACCTTTTATGAAGACTCTGAGAGCACTTCCGGGGAAGAGAACCGAGCTGGAG 1620  
DB 1561 TACATTAACCTTTTATGAAGACTCTGAGAGCACTTCCGGGGAAGAGAACCGAGCTGGAG 1620  
QY 1621 CTTGATTAAGTGGGATTTAA 1641  
DB 1621 CTTGATTAAGTGGGATTTAA 1641  
RESULT 2  
ADF47823  
ID ADF47823 standard; DNA; 1902 BP.  
XX  
AC ADF47823;  
XX

DT 26-FEB-2004 (first entry)  
XX Arabidopsis thaliana acyltransferase nucleotide sequence SEQ ID NO:8.  
XX  
XX acyltransferase; enzyme; membrane-spanning region;  
XX active membrane independent acyltransferase; fatty acid ester;  
XX fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;  
XX phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
XX acylated amino acid; structured lipid; fat-soluble molecule; gene; ds.  
OS Arabidopsis thaliana.  
XX  
XX W02003100044-A1.  
XX  
XX 04-DEC-2003.  
XX  
XX 28-MAY-2003; 2003WO-SE000870.  
XX  
XX 29-MAY-2002; 2002SE-00001581.  
XX 29-MAY-2002; 2002US-0383889P.  
XX 20-JAN-2003; 2003SE-00000142.  
XX  
XX (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.  
XX  
XX Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;  
XX  
XX WPI, 2004-053268/05.  
XX P-PSDB; ADF47824.  
XX  
XX New nucleotide sequence encoding an improved acyltransferase polypeptide,  
XX useful for producing structured lipids or fat-soluble molecules, in  
XX removing undesirable fat or in modifying lipids in animal or plant raw  
XX material.  
XX  
XX Claim 5; SEQ ID NO 8; 91pp; English.  
XX  
XX The present invention describes a nucleotide sequence (I) derived from a  
XX nucleotide sequence encoding an acyltransferase polypeptide comprising at  
XX least one membrane-spanning region. (I) encodes an improved active  
XX membrane independent acyltransferase polypeptide in which at least one  
XX amino acid residue of the membrane-spanning region has been deleted  
XX and/or substituted as compared to the original acyltransferase  
XX polypeptide, where the encoded active membrane independent  
XX acyltransferase polypeptide can produce fatty acid esters and/or fatty  
XX acid thioesters such as triacylglycerols, diacylglycerols,  
XX monoacylglycerols, phospholipids, glycolipids, lysolipids, waxesters,  
XX acylated carbohydrates and acylated amino acids. Also described: (1) a  
XX nucleotide sequence molecule (II) comprising at least one promoter region  
XX which functions in a host, where the promoter region is operably linked  
XX to at least one (I), which is operably linked to at least one non-  
XX translated region which functions in a host; (2) a vector comprising (II)  
XX; (3) a host cell comprising (II) or the vector; (4) producing an  
XX improved active membrane independent acyltransferase polypeptide,  
XX comprising providing the above host cell and a growth medium preparing a  
XX host cell culture, culturing the host cell culture and harvesting the  
XX acyltransferase polypeptide; (5) a polypeptide obtained by the method in  
XX (4), where the polypeptide at least is an improved active membrane  
XX independent acyltransferase polypeptide; (6) an oligonucleotide  
XX specifically hybridizing to (I) under stringent conditions; and (7) a kit  
XX comprising the above polypeptide and a stabiliser. The nucleic acid  
XX molecule (I) and the polypeptide are useful in producing structured  
XX lipids or fat-soluble molecules, in removing undesirable fat or in  
XX modifying lipids present in animal and plant raw material. The present  
XX sequence is used in the exemplification of the present invention.  
XX  
XX Sequence 1902 BP; 516 A; 405 C; 444 G; 537 T; 0 U; 0 Other;  
XX  
XX Query Match 98.8%; Score 1632; DB 12; Length 1902;  
XX Best Local Similarity 99.7%; Pred. No. 0;  
XX Matches 1625; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
XX  
XX 1 ATGGAGCGAATTCGAATCAGTAAACGGCTCTTCAACCGTCATCGCGTTTCTTG 60

DB  
1 ATGGAGCGAATTCGAATCAGTAAACGGCTCTTCAACCGTCATCGCGTTTCTTG 60  
QY  
61 ATTGCGGTGGCCGAATCTCGCGTGGAGATAGACCGAGTTTCACGCGACTCTCGAAG 120  
DB  
61 ATTGCGGTGGCCGAATCGCGGTGGAGATAGACCGAGTTTCACGCGACTCTCGAAG 120  
QY  
121 CTATCGGGTATTAATATTCGGGAAATTGCGTGCAGCGAGCTACGAGCGGTGCGATCTT 180  
DB  
121 CTATCGGGTATTAATATTCGGGAAATTGCGTGCAGCGAGCTACGAGCGGTGCGATCTT 180  
QY  
181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTCGACTCTGTATGCTAGACCACT 240  
DB  
181 GACTGTCCATACACTCCGTTGGACTTCAATCCGCTCGACTCTGTATGCTAGACCACT 240  
QY  
241 AAGCTTCTTCTGCTGTCACTGCTGTTTAAGTATAGTGTCTAGATCTTATATCA 300  
DB  
241 AAGCTTCTTCTGCTGTCACTGCTGTTTAAGTATAGTGTCTAGATCTTATATCA 300  
QY  
301 ACAGACCATCCCGAGTGAATGACACGGCTGACAGTGTCTTTCAGGCATCAGAAATG 360  
DB  
301 ACAGACCATCCCGAGTGAATGACACGGCTGACAGTGTCTTTCAGGCATCAGAAATG 360  
QY  
361 GATCCAGGTTACATTAACAGTCTCTTCTTCTACTGCTGGAAGAGTGGCTTAAGTGT 420  
DB  
361 GATCCAGGTTACATTAACAGTCTCTTCTTCTACTGCTGGAAGAGTGGCTTAAGTGT 420  
QY  
421 GTTGAGTTTGTATGAAAGCAATGCAATGTGCTGTTCCATACAGATTGAGATTGCA 480  
DB  
421 GTTGAGTTTGTATGAAAGCAATGCAATGTGCTGTTCCATACAGATTGAGATTGCA 480  
QY  
481 CCAACCAATGGAAGACGCTTACTTCTTCAAGTCTCAAGTGAACCTTGAACT 540  
DB  
481 CCAACCAATGGAAGACGCTTACTTCTTCAAGTCTCAAGTGAACCTTGAACT 540  
QY  
541 GCTTAAACCTCCGTGGCGCCCTTCTATAGTATTTGGCCATTCAGTGGTATATATGTC 600  
DB  
541 GCTTAAACCTCCGTGGCGCCCTTCTATAGTATTTGGCCATTCAGTGGTATATATGTC 600  
QY  
601 TTGACATCTTCTTCTGGAATGCTGAGGCTAGAAATTCACCAAAATTAATTGAAATG 660  
DB  
601 TTGACATCTTCTTCTGGAATGCTGAGGCTAGAAATTCACCAAAATTAATTGAAATG 660  
QY  
661 CTGATCGCATATCATGCTTATTTGCTGTGAGGCTCCCTCTTCTGTTCTGTTGAG 720  
DB  
661 CTGATCGCATATCATGCTTATTTGCTGTGAGGCTCCCTCTTCTGTTCTGTTGAG 720  
QY  
721 GCAATCAATCTACTCTCTCTGCTGTAAGTGTGCTTCTGTTCTGAGGAACTGCT 780  
DB  
721 GCAATCAATCTACTCTCTCTGCTGTAAGTGTGCTTCTGTTCTGAGGAACTGCT 780  
QY  
781 CGGTGTTGTCGAATCTTCTTCTGCTGTCATGTTGCTTATGCTTCAAGAAATGTC 840  
DB  
781 CGGTGTTGTCGAATCTTCTTCTGCTGTCATGTTGCTTATGCTTCAAGAAATGTC 840  
QY  
841 AAGGTGATTAACATCTCTGACCGATTTTCTGGGGTGTGCAAGAAAGTAAAGGC 900  
DB  
841 AAGGTGATTAACATCTCTGACCGATTTTCTGGGGTGTGCAAGAAAGTAAAGGC 900  
QY  
901 GTATACCACTGTGATGAGAGAGATATATCAATCAAAATATTTGCTGCGCCGCAATAT 960  
DB  
901 GTATACCACTGTGATGAGAGAGATATATCAATCAAAATATTTGCTGCGCCGCAATAT 960  
QY  
961 ATTAACATTGAATTCCTTCTCACTGACGTGTACAGAAACAGCTTATGCAATGACCGC 1020  
DB  
961 ATTAACATTGAATTCCTTCTCACTGACGTGTACAGAAACAGCTTATGCAATGACCGC 1020  
QY  
1021 ATGAATGTGGCTTCCACCGCTTCTTCTTCAACGCGCGTGAATCTAGATGGAGCT 1080  
DB  
1021 ATGAATGTGGCTTCCACCGCTTCTTCTTCAACGCGCGTGAATCTAGATGGAGCT 1080  
QY  
1081 CTTTCAAGCAATGAAGATATGACCCAGATGCAAGAGATTTACCAAGTTAAG 1140

Db	1081	CTTTTCAGAGCAATGAAAGACATACACCCAGATAGCAAGAGATCTTACACCACTTAAAG	1140
Oy	1141	AAGTTGTATCATGATGACCTCTGTTTTTAATCCTCTGACTCTTGGAGAGACCACTTAA	1200
Db	1141	AAGTTGTATCATGATGACCTCTGTTTTTAATCCTCTGACTCTTGGAGAGACCACTTAA	1200
Oy	1201	AAAAATGATTTTGGATATATATGATGCTCATCTTAAAGACAGAGGTTGGTATTACTTGGCC	1260
Db	1201	AAAAATGATTTTGGATATATGATGCTCATCTTAAAGACAGAGGTTGGTATTACTTGGCC	1260
Oy	1261	CCAAATGGCAAACTTTATCCTGATATATTTGGATCATCAGAGATATCATTTACGAAACTGAA	1320
Db	1261	CCAAATGGCAAACTTTATCCTGATATATTTGGATCATCAGAGATATCATTTAAGAAACTGAA	1320
Oy	1321	GGTTCCTCTGTCCTCAAGGCTCTGGAACCTGTGTTGATGGAACCGCTGACCTATTAACCTGG	1380
Db	1321	GGTTCCTCTGTCCTCAAGGCTCTGGAACCTGTGTTGATGGAACCGCTGACCTATTAACCTGG	1380
Oy	1381	GATGAGACGGTACCTTATCATTTCACTCTCTTGGTGCAAGAAATGGCTCGACCTTAAAGTT	1440
Db	1381	GATGAGACGGTACCTTATCATTTCACTCTCTTGGTGCAAGAAATGGCTCGACCTTAAAGTT	1440
Oy	1441	AACATTAACAATGGCTCTCCCAAGCCAGAACACCATGATGGAAACGTCATATGTGAAACTAAAT	1500
Db	1441	AACATTAACAATGGCTCTCCCAAGCCAGAACACCATGATGGAAACGTCATATGTGAAACTAAAT	1500
Oy	1501	GTTGATCATGAGCATGGGCTCAGACATCATATAGCTAACATGACAAAAGACCAAGGGTTAAG	1560
Db	1501	GTTGATCATGAGCATGGGCTCAGACATCATATAGCTAACATGACAAAAGACCAAGGGTTAAG	1560
Oy	1561	TACATTAACCTTTTATGAGAAGCTCTGAGAAGATTCGGGGGAAAGAACCGCAGTCTGGAG	1620
Db	1561	TACATTAACCTTTTATGAGAAGCTCTGAGAAGATTCGGGGGAAAGAACCGCAGTCTGGAG	1620
Oy	1621	CTTGATTAATAA	1630
Db	1621	CTTGATTAATAA	1630

RESULT 3	
AAc64438	
ID	AAc64438 standard; DNA; 3896 BP.
XX	
AC	AAc64438;
XX	
DT	12-FEB-2001 (first entry)
XX	
DE	Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:11.
XX	
KW	PDAT; Phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KW	EST; expressed sequence tag; fatty acid; oil content; ds.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200060095-A2.
XX	
PD	12-OCT-2000.
XX	
PF	28-MAR-2000; 2000WO-EP002701.
XX	
PR	01-APR-1999; 99EP-00106656.
PR	10-JUN-1999; 99EP-0011321.
PR	07-FEB-2000; 2000US-0180687P.
XX	
RA	(BADI ) BASF PLANT SCI GMBH.
XX	
PI	Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Stymer S;
XX	
DR	WPI; 2000-665012/64.
XX	
PT	Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT	pathway for triacylglycerol production and DNAs encoding them, useful for
PT	producing triacylglycerol, or for transforming any cell or organism to

PT increase oil content.  
XX  
XX Claim 6, Page 60-61; 97bp, English.  
XX  
XX The present invention describes an enzyme for catalyzing (in an acyl-CoA-independent reaction) the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol (TAG). The enzyme is designated as phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the nucleotides encoding them are useful for producing TAG and/or TAG with uncommon fatty acids. The enzyme and the nucleotide are also useful for transforming any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism. The present sequence represents the Arabidopsis thaliana PDAT genomic DNA  
XX  
XX  
SQ Sequence 3896 BP; 1042 A; 725 C; 769 G; 1360 T; 0 U; 0 Other;

Query Match 16.6%; Score 271.8; DB 3; Length 3896;  
Best Local Similarity 80.0%; Fred. No. 1.8e-74;  
Matches 375; Conservative 0; Mismatches 2; Indels 92; Gaps 1;

QY	768	TGAGGAATCTGCTGGTGTGTGTCCAATCTTTTGGCGCGCATTTGAGGCTTAGGCATT	827
Db	1787	TCAGGAACTGCTCGTGTGTGTGTCCAATCTTTTGGCGCGCATTTGAGGCTTAGGCATT	1846
QY	828	TTCAAGAATTGCAAAGGGTGATTAACACATCTGAGCGCATTTTCTGGGGGTGCTGCAAA	887
Db	1847	TTCAAGAATTGCAAAGGGTGATTAACACATCTGAGCGCATTTTCTGGGGGTGCTGCAAA	1906
QY	888	GAAAGATATAGGCGCGATATCCACTGTGTGATGAAAGAGAAATATCAAAATTTCTGGCTG	947
Db	1907	GAAAGATATAGGCGCGATATCCACTGTGTGATGAAAGAGAAATATCAAAATTTCTGGCTG	1966
QY	948	GCCGCAAAATTTATTTAACTTTAACTTTCTTCCACTAGC-----	987
Db	1967	GCCGCAAAATTTATTTAACTTTAACTTTCTTCCACTAGGCGTATAGACTGTGTATATGCA	2026
QY	988	-----	987
Db	2027	ACTGTAACTATACAAAGTTTCAACAAGATGTTCATCTCATATTTGTTCTCTTGAT	2086
QY	988	-----GTTACAGAAACGCTCTACTCAAACTGACACAGATGGATGTGGCCTT	1039
Db	2087	GTTGATTCATAGTTATACAGAAACGCTCTACTCAAACTGACACAGATGGATGTGGCCTT	2146
QY	1036	CCCAACCCCTTTGTCTTTTCAACAGCCCGTGAATCAGATGGAGCTCTTTTCAAAACATA	1099
Db	2147	CCCAACCCCTTTGTCTTTTCAACAGCCCGTGAATCAGATGGAGCTCTTTTCAAAACATA	2206
QY	1096	GAAGACTATGACCCGATAGCAAGAGATGTTTACACCACTTAAAGAGT	1144
Db	2207	GAAGACTATGACCCGATAGCAAGAGATGTTTACACCACTTAAAGAGT	2255

	RESULT 4
AC64448	AAC64448 standard; DNA, 3896 BP.
ID	AAC64448
XX	
AC	AAC64448;
XX	
DT	12-FEB-2001 (first entry)
XX	
DE	Arabidopsis thaliana PDAT genomic DNA SEQ ID NO:10b.
XX	
KW	PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KW	EST; expressed sequence tag; fatty acid; oil content; ds.
XX	
OS	Arabidopsis thaliana.
XX	
PN	WO200060095-A2.
XX	
PD	12-OCT-2000.

```

XX 28-MAR-2000; 2000MO-EP002701.
PF 01-APR-1999; 99EP-00106656.
XX 10-JUN-1999; 99EP-00113321.
PR 07-FEB-2000; 2000US-0180687P.
XX (BADI ) BASF PLANT SCI GMBH.
XX Dahlqvist A, Stahl U, Lemman M, Banas A, Ronne H, Styenne S;
PI WPI; 2000-665012/64.
XX
XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DNAs encoding them, useful for
PT producing triacylglycerol, or for transforming any cell or organism to
PT increase oil content.
XX
XX Claim 6; Page 95-96; 97pp; English.
XX
XX The present invention describes an enzyme for catalysing (in an acyl-CoA-
CC independent reaction) the transfer of fatty acids from phospholipids to
CC diacylglycerol in the biosynthetic pathway for the production of
CC triacylglycerol (TAG). The enzyme is designated as
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the
CC nucleotides encoding them are useful for producing TAG and/or TAG with
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for
CC transforming any cell or organism in order to be expressed in this cell
CC or organism and result in an altered, preferably increased oil content of
CC this cell or organism. The present sequence represents the Arabidopsis
CC thaliana PDAT genomic DNA
XX
SQ Sequence 3896 BP; 1042 A; 733 C; 760 G; 1360 T; 0 U; 1 Other;
Query Match 16.6%; Score 271.8; DB 3; Length 3896;
Best Local Similarity 80.0%; Pred. No. 1.8e-74;
Matches 375; Conservative 0; Mismatches 2; Indels 92; Gaps 1;
QY 768 TGAGGAACTGCTGGTGTGTTGTCATTTTGGTGTGATGCTTATGCCATT 827
DB 1787 TCAGGAACTGCTGGTGTGTTGTCATTTTGGTGTGATGCTTATGCCATT 1846
QY 828 TTCAAGAATTSCAAGGATGATATACATCTGAGAGCATTTTCTGGGGGAGTGCAGAA 887
DB 1847 TTCAAGAATTSCAAGGATGATATACATCTGAGAGCATTTTCTGGGGGAGTGCAGAA 1906
QY 888 GAAAGATTAAGCGGATATACACTGTGATGAGAGGAATATCAATCAATAATTCTGGCTG 947
DB 1907 GAAAGATTAAGCGGATATACACTGTGATGAGAGGAATATCAATCAATAATTCTGGCTG 1966
QY 948 GCGGCAATATTTATTAACATTGAATTCTTCCACTAGC----- 987
DB 1967 GCGGCAATATTTATTAACATTGAATTCTTCCACTAGC----- 987
QY 988 ----- 987
DB 2021 ACTGTAACCTAACAAAGTTTCCAGAAAGTTTCACTCTATATTTGTTCTTTGAT 2086
QY 988 -----GTTACAGAAAAGCTCTAGTCAACATGACGAGCATGAGATGTGGCTT 1035
DB 2087 GTGTATTCATCAGTTTACAGAAAAGCTCTAGTCAACATGACGAGCATGAGATGTGGCTT 2146
QY 1036 CCCACCCCTTTTGTCTTTTCACAGCCCGTGAATGACAGATGGAGCTTTTTCAGAGCAATA 1095
DB 2147 CCCACCCCTTTTGTCTTTTCACAGCCCGTGAATGACAGATGGAGCTTTTTCAGAGCAATA 2206
QY 1096 GAAAGTATGACCCAGATAGCAAGAGATGTTTACACCACTTAAAGAAAT 1144
DB 2207 GAAAGTATGACCCAGATAGCAAGAGATGTTTACACCACTTAAAGAAAT 2255

```

RESULT 5  
AAS01096

```

ID AAS01096 standard; cDNA; 328 BP.
XX
XX AAS01096;
AC 31-MAY-2001 (first entry)
XX
XX Corn sterol acyltransferase LCAT EST sequence #3.
DE
XX Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;
XX acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;
XX nutritional supplement; dairy product; food product; salad dressing;
XX corn; Arabidopsis thaliana; expressed sequence tag; EST; ss.
XX
XX Zea mays.
OS
XX WO200116308-A2.
XX
XX 08-MAR-2001.
XX
XX 30-AUG-2000; 2000MO-US023863.
XX
XX 30-AUG-1999; 99US-0152493P.
XX
XX (MONS ) MONSANTO CO.
XX
XX Lasener M, Van Benenham A;
PI WPI; 2001-169010/17.
XX
XX New isolated nucleic acid encoding plant lecithin:cholesterol
PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase
PT -like polypeptides, for modifying the sterol content and oil production
PT of plants.
XX
XX Claim 5; Page 90; 127pp; English.
XX
XX The present sequence for corn LCAT (lecithin:cholesterol acyltransferase-
CC like) EST sequence #3 is closely related to the Arabidopsis thaliana
CC LCAT2 sequence. Several novel polynucleotides encoding the plant sterol
CC acyltransferases LCAT (AAS001081-AAS01104, AAS01341) and ACAT (acyl
CC CoA:cholesterol acyltransferase-like; AAS01311-AAS01319) are described. A
CC yeast LCAT related open reading frame, LRO1 gene sequence (AAS01342), and
CC a rat ACAT (AAS01105) cDNA sequence are also described. The
CC polynucleotides encoding LCAT or ACAT are used to produce LCAT or ACAT
CC polypeptides. They can also be used in a recombinant construct to
CC transform a host cell (preferably of a plant) or a plant. The recombinant
CC construct is used to increase or decrease the sterol content of the host
CC cell or plant. It can be used to alter oil production of the cell or
CC plant, preferably by increasing it. The oil of the plant or the plant
CC itself is used as a food product, or as nutritional or dietary
CC supplements, or in pharmaceutical compositions for lowering cholesterol.
CC The oil can be used in foods e.g. margarine, butter, cooking oil, and
CC dressings e.g. salad dressings, mayonnaise, cheese, processed meat,
CC pasta, sauces, cereals, desserts, dips, chips, baked goods, pastries,
CC cookies, snack bars, confections, chocolates, and beverages. The
CC alteration in sterol content and/or composition can also provide a plant
CC with tolerance to stress and insect damage
XX
SQ Sequence 328 BP; 86 A; 68 C; 73 G; 101 T; 0 U; 0 Other;
Query Match 10.7%; Score 176.2; DB 5; Length 328;
Best Local Similarity 79.0%; Pred. No. 9.2e-45;
Matches 222; Conservative 0; Mismatches 58; Indels 1; Gaps 1;
QY 228 ACTAAGCTTCTTCTGCTGCTCACTGCTGTTAAGTATGCTGATGCTTATAT 297
DB 35 ACATGCTTTTCTGCGGTAAATGCTGCTTAATGATCAAGCTCTGAACCCCTATAT 94
QY 298 GAAAGACCAATCCGAGATTAAGTCAAGGCGCTGACATGCTCTTTCAGGCATCAGAA 357
DB 95 CAGATGACCAATCCGAGATTAAGTCAAGGCGCTGATGATGCTC-TTTCGCAATTTACAGG 153
QY 358 TTGATTCAGGTTACATTAACAGTCTCTTCTACTGTCGTGAAAGAGTGCTTAAGTG 417

```

Db 154 CTGAGCCCTGGTATATATAACAGCTCTCTCTTGAGTATGAAAGATGGCTCAATGG 213  
Oy 418 TGTGTGAGTTGGTATAGAACCAATGCAATTTGCGCTTCCATAGCATTGGAGATTG 477  
Db 214 TGTGTAGAGTTGGGATTTGAACTGAACTAATGCAATTTATCGCTTCCGTATGATTGAGACTG 273  
Oy 478 TCACCAACCAATTTGGAAGAGCGTACCTTTACTTTACCA 518  
Db 274 CCCCATCAATGCTTGAGAGAGAGATCTGTACTTTCA 314

## RESULT 6

AAC64449  
ID AAC64449 standard; cDNA; 709 BP.

XX AAC64449;

XX 12-FEB-2001 (first entry)

XX Lycopersicon esculentum PDAT nucleotide sequence SEQ ID NO:11b.

XX PDAT, phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;  
KM EST; expressed sequence tag; fatty acid; oil content; ss.

XX Lycopersicon esculentum.

XX MO200060095-A2.

XX 12-OCT-2000.

XX 28-MAR-2000; 2000WO-EP002701.

XX 01-APR-1999; 99EP-00106656.

XX 10-JUN-1999; 99EP-00111321.

XX 07-FEB-2000; 2000US-0180687P.

XX (BADI ) BASF PLANT SCI GMBH.

XX Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Skymne S;

XX WPI; 2000-665012/64.

XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic  
PT pathway for triacylglycerol production and DNAs encoding them, useful for  
PT producing triacylglycerol, or for transforming any cell or organism to  
PT increase oil content.

XX Claim 6; Page 97; 97BP; English.

XX The present invention describes an enzyme for catalysing (in an acyl-CoA-  
CC independent reaction) the transfer of fatty acids from phospholipids to  
CC diacylglycerol in the biosynthetic pathway for the production of  
CC triacylglycerol (TAG). The enzyme is designated as

CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the  
CC nucleotides encoding them are useful for producing TAG and/or TAG with  
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for  
CC transforming any cell or organism in order to be expressed in this cell  
CC or organism and result in an altered, preferably increased oil content of  
CC this cell or organism. The present sequence represents the Lycopersicon  
CC esculentum PDAT nucleotide sequence

XX Sequence 709 BP; 226 A; 121 C; 161 G; 201 T; 0 U; 0 Other;

XX Query Match 6.7%; Score 110.4; DB 3; Length 709;

XX Best Local Similarity 74.5%; Pred. No. 1.1e-23;

XX Matches 152; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

Oy 1426 CTCGAGCTTAAGTTAACTAATGAGCTCCCGAGCCAGAACACAGATGAGAGCGTA 1485

Db 1 CTGGGGCCAAAAGTAACATTAACAAGACACACAGCTGAGACATGAT-GTTCAATGTA 59

Oy 1486 CATGTGAACCTAAATGTTGATCATGAGCATGGGTCAACATCATATGCTTAACATGACAAA 1545

Db 60 CAAGTGCATCTTAATATAGACATCAACATGTTGAAGATATCATTTCCCAATATGACAAAG 119  
Oy 1546 GCACCAAGGTTTAACTAATACCTTTTATGAAGACTTGAGAGCATTTCCGGGAAAGGA 1605  
Db 120 TTACTTAAAGTAAGTAATACATTAACCTATTATGAGATTCTGAAAGTTTCCAGGACAAAG 179  
Oy 1606 ACCGAGTCTGGAGCTTGATTA 1629  
Db 180 ACAGCATTTGGAGCTTGATTA 203

## RESULT 7

AAC64439  
ID AAC64439 standard; cDNA; 709 BP.

XX AAC64439;

XX 12-FEB-2001 (first entry)

XX Lycopersicon esculentum PDAT nucleotide sequence SEQ ID NO:12.

XX PDAT, phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;  
KM EST; expressed sequence tag; fatty acid; oil content; ss.

XX Lycopersicon esculentum.

XX MO200060095-A2.

XX 12-OCT-2000.

XX 28-MAR-2000; 2000WO-EP002701.

XX 01-APR-1999; 99EP-00106656.

XX 10-JUN-1999; 99EP-00111321.

XX 07-FEB-2000; 2000US-0180687P.

XX (BADI ) BASF PLANT SCI GMBH.

XX Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Skymne S;

XX WPI; 2000-665012/64.

XX Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic  
PT pathway for triacylglycerol production and DNAs encoding them, useful for  
PT producing triacylglycerol, or for transforming any cell or organism to  
PT increase oil content.

XX Claim 6; Page 62; 97BP; English.

XX The present invention describes an enzyme for catalysing (in an acyl-CoA-  
CC independent reaction) the transfer of fatty acids from phospholipids to  
CC diacylglycerol in the biosynthetic pathway for the production of  
CC triacylglycerol (TAG). The enzyme is designated as

CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the  
CC nucleotides encoding them are useful for producing TAG and/or TAG with  
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for  
CC transforming any cell or organism in order to be expressed in this cell  
CC or organism and result in an altered, preferably increased oil content of  
CC this cell or organism. The present sequence represents the Lycopersicon  
CC esculentum PDAT nucleotide sequence

XX Sequence 709 BP; 226 A; 120 C; 162 G; 201 T; 0 U; 0 Other;

XX Query Match 6.7%; Score 110.4; DB 3; Length 709;

XX Best Local Similarity 74.5%; Pred. No. 1.1e-23;

XX Matches 152; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

Oy 1426 CTCGAGCTTAAGTTAACTAATGAGCTCCCGAGCCAGAACACAGATGAGAGCGTA 1485

Db 1 CTGGGGCCAAAAGTAACATTAACAAGACACACAGCTGAGACATGAT-GTTCAATGTA 59

Oy 1486 CATGTGAACCTAAATGTTGATCATGAGCATGGGTCAACATCATATGCTTAACATGACAAA 1545

Db 60 CAAGTGCATCTTAATATAGACATCAACGTGTGAAGATATCTTCCCAATATGCAAAAG 119  
Qy 1546 GCACCAAGGGTTAAAGTACATAACCTTTATGGAAGACTGTGAGACATTCGGGGAGAGCA 1605  
Db 120 TTACTTACAAATGAAATACATAACCTATATATGAGATTCTGAAAGATTTCACGGGACAAAGA 179  
Qy 1606 ACCGAGTGTGGAGCTTGATATA 1629  
Db 180 ACAGCAGTTTGGAGCTTGATATA 203

RESULT 8  
AAS01097  
ID AAS01097 standard; cDNA; 356 BP.  
AC AAS01097;  
DT 31-MAY-2001 (first entry)  
XX  
XX  
DE Corn sterol acyltransferase LCAT EST sequence #4.  
XX  
XX lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;  
KM acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;  
KM nutritional supplement; dairy product; food product; salad dressing;  
KM corn; Arabidopsis thaliana; expressed sequence tag; EST; ss.  
XX  
OS Zea mays.  
XX  
XX WO200116308-A2.  
XX  
XX 08-MAR-2001.  
XX  
XX 30-AUG-2000; 2000WO-US023863.  
XX  
XX PR 30-AUG-1999; 99US-0152493P.  
XX  
XX (MONS ) MONSANTO CO.  
XX  
XX Laesener M, Van Eenennaam A;  
XX  
XX WPI; 2001-169010/17.  
XX  
XX  
XX New isolated nucleic acid encoding plant lecithin:cholesterol  
PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase  
PT -like polypeptides, for modifying the sterol content and oil production  
PT of plants.  
XX  
PS Claim 5; Page 90; 127pp; English.

The present sequence for corn LCAT (lecithin:cholesterol acyltransferase-like) EST sequence #4 is closely related to the Arabidopsis thaliana LCAT2 sequence. Several novel polynucleotides encoding the plant sterol acyltransferase LCAT (AAS001081-AAS01104, AAS01341) and ACAT (acyl CoA:cholesterol acyltransferase-like; AAS01311-AAS01319) are described. A yeast LCAT related open reading frame, LK01 gene sequence (AAS01342), and a rat ACAT (AAS01105) cDNA sequence are also described. The polynucleotides encoding LCAT or ACAT are used to produce LCAT or ACAT polypeptides. They can also be used in a recombinant construct to transform a host cell (preferably of a plant) or a plant. The recombinant construct is used to increase or decrease the sterol content of the host cell or plant. It can be used to alter oil production of the cell or plant, preferably by increasing it. The oil of the plant or the plant itself is used as a food product, or as nutritional or dietary supplements, or in pharmaceutical compositions for lowering cholesterol. The oil can be used in foods e.g. margarine, butter, cooking oil, and dressings e.g. salad dressings, mayonnaise, cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips, baked goods, pastries, cookies, snack bars, confections, chocolates, and beverages. The alteration in sterol content and/or composition can also provide a plant with tolerance to stress and insect damage

Sequence 356 BP; 98 A; 80 G; 78 G; 100 T; 0 U; 0 Other;

Query Match 6.6%; Score 109; DB 5; Length 356;  
Best Local Similarity 76.1%; Pred. No. 2e-23;  
Matches 150; Conservative 0; Mismatches 40; Indels 7; Gaps 1;

Qy 338 GTCTTCAGCCATCAGAAATTTGATCCAGGTTACATTA-----CAGTCCCTTTCT 390  
Db 1 GTCTTCAGCAATTAACAGCTGACCTGGTTATATACAGGTTTCAAGTCTCTCTCT 60  
Qy 391 ACTGTCGAAAGAGTGGCTTAAGTGTGTGTGTGAGTTTGGTATATGAAGCAATTCATTT 450  
Db 61 TCAGTATGAAAGAAATGGATGGATCAATGGTGTAGAGTTTGCAATGAAGCTTAATCAATT 120  
Qy 451 GTCCGTGTTCCATTCGATTTGAGATTTGTCACCAACCAATTTGSAAGAGCGTACCTTAC 510  
Db 121 ATCGCTGTCCGTATGATGAGAGACTGCCCATCAATGCTTGAAGAGAGAGATCTGTAC 180  
Qy 511 TTTCACAGCTCAAGTT 527  
Db 181 TTTCACAAATTAAGTT 197

RESULT 9  
ADF47831  
ID ADF47831 standard; DNA; 1872 BP.  
AC ADF47831;  
XX  
XX  
XX 26-FEB-2004 (first entry)  
XX  
XX  
DE Schizosaccharomyces pombe acyltransferase DNA sequence SEQ ID NO:16.  
XX  
XX  
XX acyltransferase; enzyme; membrane-spanning region;  
KM active membrane independent acyltransferase; fatty acid ester;  
KM fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;  
KM phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
KM acylated amino acid; structured lipid; fat-soluble molecule; gene; ds.  
XX  
OS Schizosaccharomyces pombe.  
XX  
XX WO2003100044-A1.  
XX  
XX 04-DEC-2003.  
XX  
XX 28-MAY-2003; 2003WO-SB000870.  
XX  
XX PR 29-MAY-2002; 2002SE-00001581.  
XX  
XX PR 29-MAY-2002; 2002US-0383889P.  
XX  
XX PR 20-JAN-2003; 2003SE-00000142.  
XX  
XX (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.  
XX  
XX Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;  
XX  
XX WPI; 2004-053268/05.  
XX  
XX P-PSDB; ADF47832.  
XX  
XX New nucleotide sequence encoding an improved acyltransferase polypeptide,  
PT useful for producing structured lipids or fat-soluble molecules, in  
PT removing undesirable fat or in modifying lipids in animal or plant raw  
PT material.  
XX  
PS Claim 5; SEQ ID NO 16; 91pp; English.

The present invention describes a nucleotide sequence (I) derived from a nucleotide sequence encoding an acyltransferase polypeptide comprising at least one membrane-spanning region. (I) encodes an improved active membrane independent acyltransferase polypeptide in which at least one amino acid residue of the membrane-spanning region has been deleted and/or substituted as compared to the original acyltransferase polypeptide, where the encoded active membrane independent acyltransferase polypeptide can produce fatty acid esters and/or fatty acid thioesters such as triacylglycerols, diacylglycerols,



CC monoacylglycerols, phospholipids, glycolipids, lysolipids, waxes, waxes, waxes, acylated carbohydrates and acylated amino acids. Also described: (1) a nucleotide sequence molecule (ii) comprising at least one promoter region CC which functions in a host, where the promoter region is operably linked CC to at least one (I), which is operably linked to at least one non- CC translated region which functions in a host; (2) a vector comprising (II) CC / (3) a host cell comprising (II) or the vector; (4) producing an CC improved active membrane independent acyltransferase polypeptide, CC comprising providing the above host cell and a growth medium preparing a CC host cell culture, culturing the host cell culture and harvesting the CC host cell culture and recovering the improved active membrane independent CC acyltransferase polypeptide; (5) a polypeptide obtained by the method in CC (4), where the polypeptide at least is an improved active membrane CC independent acyltransferase polypeptide; (6) an oligonucleotide CC specifically hybridizing to (I) under stringent conditions; and (7) a kit CC comprising the above polypeptide and a stabiliser. The nucleic acid CC molecule (I) and the polypeptide are useful in producing structured CC lipids or fat-soluble molecules, in removing undesirable fat or in CC modifying lipids present in animal and plant raw material. The present CC sequence is used in the exemplification of the present invention.

XX Sequence 1872 BP; 598 A; 297 C; 405 G; 572 T; 0 U; 0 Other;

Query Match 3.2%; Score 52.6; DB 12; Length 1872;  
Best Local Similarity 47.4%; Pred. No. 4.1e-05;  
Matches 157; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

```

Qy 420 TGTGAGTTGGTATGAAAGCAATTCGCTGTCATCGATTGAGATTGTC 479
Db 714 TGTCTCAATTTGTTATGAGCCATTAACATGTTAAAGTCTTTACGATTGGCGTTATC 773
Qy 480 ACCAACCAATTTGAAAGCGCTGACCTTACTTTCACAGCTCAAGTTGACCTTTGAAC 539
Db 774 ATATCAATATTAGAGAAACGATTAATATTTTCAAAAGTTAAAGTTGATGAGTA 833
Qy 540 TGCCTTAAACCTCCGTCGCGCCCTTCTATAGTATTTGGCCATTCAATGGTTAATAT 599
Db 834 CAGCAACATTTGATCAATGAGAAAGATGTTGATTTCTCACTCCATGGGTTCAAGGT 893
Qy 600 CTTGACATATCTTTCGAATGCGCTGAGCTAGAAATTTGACCAAAACATTAATTGAGTC 659
Db 894 TACGACTATTTTAAAGTGGTTGAAGCTAGGCTACGAAATGCTGACCACTTG 953
Qy 660 GCTTATGACATATTCATGCTTATTTGCTGCTGTTGAGCTCCTCTTCTTGTCTGTTGA 719
Db 954 GATTATATGATCATATTGAAGCATTTATAATATCGGATCTTTGATTGAGACCCAA 1013
Qy 720 GGCATCAATTAATCTCTCTCTGCTGTTAAG 750
Db 1014 AACAGTGGCAGCGCTTTTATCGGGTGAATG 1044

```

RESULT 10

ADFA7817 ADFA7817 standard; DNA; 1701 BP.

ADFA7817;

26-FEB-2004 (first entry)

Yeast membrane independent acyltransferase DNA sequence SEQ ID NO:2.

XX acyltransferase; enzyme; membrane-spanning region;  
XX active membrane independent acyltransferase; fatty acid ester;  
XX fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;  
XX phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
XX acylated amino acid; structured lipid; fat-soluble molecule; yeast; gene;  
XX ds.

XX Saccharomyces cerevisiae.

XX MO2003100044-A1.

PD 04-DEC-2003.

XX 28-MAY-2003; 2003MO-SE000870.

XX 29-MAY-2002; 2002SE-00001581.

XX 29-MAY-2002; 2002US-038389P.

XX 20-JAN-2003; 2003SE-00000142.

XX (SCAN) SCANDINAVIAN BIOTECHNOLOGY RES AB.

XX Dahlqvist A, Ghosal A, Lindqvist Y, Banae A;

XX WPI; 2004-053268/05.

PT New nucleotide sequence encoding an improved acyltransferase polypeptide, useful for producing structured lipids or fat-soluble molecules, in removing undesirable fat or in modifying lipids in animal or plant raw material.

PS Claim 12; SEQ ID NO 2; 91pp; English.

XX The present invention describes a nucleotide sequence (I) derived from a CC nucleotide sequence encoding an acyltransferase polypeptide comprising at CC least one membrane-spanning region. (I) encodes an improved active CC membrane independent acyltransferase polypeptide in which at least one CC amino acid residue of the membrane-spanning region has been deleted CC and/or substituted as compared to the original acyltransferase CC polypeptide, where the encoded active membrane independent CC acyltransferase polypeptide can produce fatty acid esters and/or fatty CC acid thioesters such as triacylglycerols, diacylglycerols, CC monoacylglycerols, phospholipids, glycolipids, lysolipids, waxes, waxes, waxes, acylated carbohydrates and acylated amino acids. Also described: (1) a CC nucleotide sequence molecule (II) comprising at least one promoter region CC which functions in a host, where the promoter region is operably linked CC to at least one (I), which is operably linked to at least one non- CC translated region which functions in a host; (2) a vector comprising (II) CC / (3) a host cell comprising (II) or the vector; (4) producing an CC improved active membrane independent acyltransferase polypeptide, CC comprising providing the above host cell and a growth medium preparing a CC host cell culture, culturing the host cell culture and harvesting the CC host cell culture and recovering the improved active membrane independent CC acyltransferase polypeptide; (5) a polypeptide obtained by the method in CC (4), where the polypeptide at least is an improved active membrane CC independent acyltransferase polypeptide; (6) an oligonucleotide CC specifically hybridizing to (I) under stringent conditions; and (7) a kit CC comprising the above polypeptide and a stabiliser. The nucleic acid CC molecule (I) and the polypeptide are useful in producing structured CC lipids or fat-soluble molecules, in removing undesirable fat or in CC modifying lipids present in animal and plant raw material. The present CC sequence is used in the exemplification of the present invention.

XX Sequence 1701 BP; 516 A; 317 C; 401 G; 467 T; 0 U; 0 Other;

Query Match 3.0%; Score 49.4; DB 12; Length 1701;  
Best Local Similarity 49.2%; Pred. No. 0.0004;  
Matches 162; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

```

Qy 428 TTGGTATGAGCAATGATTCGCTGTCATCGATTGAGATTGACCAACCA 487
Db 530 TTGGCTATGAAACCCATTAATATGACAGTCTGGTATGATGAGGCTTGCAATTTAG 589
Qy 488 AATTGAAAGCGCTGACCTTACTTTCACAGCTCAAGTGAACCTTGAACCTTTAA 547
Db 590 ATCTAGAAAGACGGGATAGTACTTTACGAGCTTAAGAAACAATGAACTGTTTATC 649
Qy 548 AACTCCGTCGCGCCCTTCTATAGTATTTGCCCATTCATGGGTTAATATCTTCAGAT 607
Db 650 AATTGAGGCTGAAAGAAAGTTGTTAATTTGACATTCATGGGTTCTCAGATTATCTTT 709
Qy 608 ACTTTCGAGATGGCTGAGCTAGAAATTTGACCA-----AAACATTAATTGAAGTGC 661
Db 710 ACTTATGAAATGGGTGAGGCTGAAAGGCTCTTTACGGTTAATGTTGCTGCTGGG 769

```



```

OY 662 TTGATCAGCATATCCATGCTTATTTGCGTGTGAGAGCTCTCTCTGTTGTTGAGG 721
DB 770 TTAACGAAACATAGATTCATTATTCATTAATGACAGGAGCGCTTGGGCGCTCCAAAG 829
OY 722 CAATCAATCTACTCTCTCTGTTGTTAACG 750
DB 830 CAGTTCAGCTCTTAATTAGTGGTGAATG 858

RESULT 11
AAC64431
ID AAC64431 standard; DNA; 1986 BP.
XX
AC AAC64431;
XX
DT 12-FEB-2001 (first entry)
XX
DE Saccharomyces cerevisiae PDAT gene SEQ ID NO:1.
XX
KW PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KM EST; expressed sequence tag; fatty acid; oil content; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200060095-A2.
XX
PD 12-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-EP002701.
XX
PR 01-APR-1999; 99EP-00106656.
PR 10-JUN-1999; 99EP-00111321.
PR 07-FEB-2000; 2000US-0180687P.
XX
PA (BADI ) BASF PLANT SCI GMBH.
XX
PI Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Styhme S;
XX
DR WPI; 2000-665012/64.
XX
DR P-PSDB; AAB24256.
XX
PT Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DNAs encoding them, useful for
PT producing triacylglycerol, or for transforming any cell or organism to
PT increase oil content.
XX
PS Claim 6; Page 45-48; 97pp; English.
XX
CC The present invention describes an enzyme for catalysing (in an acyl-CoA-
CC independent reaction) the transfer of fatty acids from phospholipids to
CC diacylglycerol in the biosynthetic pathway for the production of
CC triacylglycerol (TAG). The enzyme is designated as
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the
CC nucleotides encoding them are useful for producing TAG and/or TAG with
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for
CC transforming any cell or organism in order to be expressed in this cell
CC or organism and result in an altered, preferably increased oil content of
CC this cell or organism. The present sequence encodes yeast (Saccharomyces
CC cerevisiae) PDAT
XX
SQ Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;

Query Match 3.0%; Score 49.4; DB 3; Length 1986;
Best Local Similarity 49.2%; Pred. No. 0.00044;
Matches 162; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

OY 428 TTGCTATGAGCAATGCAATGCTGCTGTCATACGATTGAGATTGTCACCAACA 487
DB 815 TTGGCTATGACCAATGCAATGCAATGCTGCTGTCATACGATTGAGATTGTCACCAACA 874
OY 488 AATTGAGAGCGTGAAGCTTACTTACCTTACCAAGCTCAAGTGAAGCTTGAAGCTTTAA 547
DB 875 ATCTAGAAAAGACGCGATAGTACTTACCAAGCTCAAGTGAAGCAATGCAAGCTTTTCAATC 934

```

```

OY 548 AACTCGTCGCGGCCCTTCTATAGTATTTGCCCATTCAATGGGTAAATATGCTTCAGAT 607
DB 935 AATTGAGTGGTGAATAAAGTTGTTTAATTGACATTTATGGGTTCTCAGATTATCTTTT 994
OY 608 ACTTTCGGAATGCGTGAAGGCTAGAAATGACCA-----AAACATTATTTGAAGTGC 661
DB 995 ACTTTATGAAATGGTCAAGGCTGAAGGCCCTCTTACGATTAATGATGATCGTGGCTGGG 1054
OY 662 TTGATCAGCATATCCATGCTTATTTGCGTGTGAGAGCTCTCTCTGTTGTTGAGG 721
DB 1055 TTAACGAAACATAGATTCATTATTAATGACAGGAGCGCTTGGGCGCTCCAAAG 1114
OY 722 CAATCAATCTACTCTCTCTGTTGTTAACG 750
DB 1115 CAGTTCAGCTCTTAATTAGTGGTGAATG 1143

RESULT 12
AAC64441
ID AAC64441 standard; DNA; 1986 BP.
XX
AC AAC64441;
XX
DT 12-FEB-2001 (first entry)
XX
DE Saccharomyces cerevisiae PDAT gene SEQ ID NO:1b.
XX
KW PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG;
KM EST; expressed sequence tag; fatty acid; oil content; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200060095-A2.
XX
PD 12-OCT-2000.
XX
PF 28-MAR-2000; 2000WO-EP002701.
XX
PR 01-APR-1999; 99EP-00106656.
PR 10-JUN-1999; 99EP-00111321.
PR 07-FEB-2000; 2000US-0180687P.
XX
PA (BADI ) BASF PLANT SCI GMBH.
XX
PI Dahlqvist A, Stahl U, Lenman M, Banas A, Ronne H, Styhme S;
XX
DR WPI; 2000-665012/64.
XX
DR P-PSDB; AAB24266.
XX
PT Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic
PT pathway for triacylglycerol production and DNAs encoding them, useful for
PT producing triacylglycerol, or for transforming any cell or organism to
PT increase oil content.
XX
PS Claim 6; Page 77-81; 97pp; English.
XX
CC The present invention describes an enzyme for catalysing (in an acyl-CoA-
CC independent reaction) the transfer of fatty acids from phospholipids to
CC diacylglycerol in the biosynthetic pathway for the production of
CC triacylglycerol (TAG). The enzyme is designated as
CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the
CC nucleotides encoding them are useful for producing TAG and/or TAG with
CC uncommon fatty acids. The enzyme and the nucleotide are also useful for
CC transforming any cell or organism in order to be expressed in this cell
CC or organism and result in an altered, preferably increased oil content of
CC this cell or organism. The present sequence represents the yeast
CC (Saccharomyces cerevisiae) PDAT gene
XX
SQ Sequence 1986 BP; 613 A; 366 C; 479 G; 528 T; 0 U; 0 Other;

Query Match 3.0%; Score 49.4; DB 3; Length 1986;
Best Local Similarity 49.2%; Pred. No. 0.00044;

```

Matches 162; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

QY 428 TTGGTATGAAAGCAATGCAATGTCGCTGTCATAGCATTTGGAGATTGTCACCAACA 487  
 |||||  
 Db 815 TTGGCTATGAAACCCATTAATAATGACGAGTGTGCTGATATTTGAGGCTTGCAATTTAG 874  
 488 AATTGGAAGAGCGTGAACCTTTACTTTCACAGAGCTCAAGTTGACCTTTGAACTGCTTTAA 547  
 |||||  
 Db 875 ATCTGAAAAGACGCATAGTACTTTACAGAGCTAAAGAACAAATCGAACTGTTTCATC 934  
 548 AACTCCGTCGCGCCCTTCTATAGTATTTGCCATTCATATGGGTATATATGCTTCAGAT 607  
 |||||  
 Db 935 AATTGAGTGTGAAAAAGTTGTTTATTTGACATTTCTATGAGTCTCAATTAATCTTTT 994  
 608 ACTTCTGGAATGCTGAGGCTAGAAATTCACCA-----AAACATTAATTTGAAAGTGC 661  
 |||||  
 QY 995 ACTTATGAAATGGGTGAGGCTGAGAGGCCCTTTACGCTAATGCTGCTGCTGAG 1054  
 662 TTGATCAGCATATCCATGCTTAATTCGCTGTTGAGCTCTCTTCTTGGTCTGTTGAG 721  
 |||||  
 Db 1055 TTAAAGAAACATAGACTCATTCATTAAATGACAGAGGAGGCTTGGGCGCTCCAAAG 1114  
 722 CAATCAATCTACTCTCTCTGCTGTAACG 750  
 |||||  
 QY 1115 CAGTTCAGCTTACTTAAGTGTGAATG 1143

## RESULT 13

AAC64440  
 ID AAC64440 standard; DNA; 1986 BP.

AC AAC64440;

DT 12-FEB-2001 (first entry)

DE Saccharomyces cerevisiae PDAT ORF nucleotide sequence SEQ ID NO:4a.

KW PDAT; phospholipid:diacylglycerol acyltransferase; triacylglycerol; TAG; EST; expressed sequence tag; fatty acid; oil content; ds.

OS Saccharomyces cerevisiae.

PN WO20006095-A2.

PD 12-OCT-2000.

PF 28-MAR-2000; 2000WO-EP002701.

PR 01-APR-1999; 99EP-00106656.

PR 10-JUN-1999; 99EP-00111321.

PR 07-FEB-2000; 2000US-0180687P.

PA (BAD1 ) BASF PLANT SCI GMBH.

PI Dahlqvist A, Stahl U, Lemman M, Banas A, Ronne H, Stryme S;

DR WPI; 2000-665012/64.

DR P-PSDB; AAB24265.

PT Phospholipid:diacylglycerol acyltransferase enzymes in the biosynthetic

PS pathway for triacylglycerol production and DNAs encoding them, useful for

XX increasing oil content.

XX Claim 6; Page 71-74; 97pp; English.

XX The present invention describes an enzyme for catalysing (in an acyl-CoA-

CC independent reaction) the transfer of fatty acids from phospholipids to

CC diacylglycerol in the biosynthetic pathway for the production of

CC triacylglycerol (TAG). The enzyme is designated as

CC phospholipid:diacylglycerol acyltransferase (PDAT). The enzyme and the

CC nucleotides encoding them are useful for producing TAG and/or TAG with

CC uncommon fatty acids. The enzyme and the nucleotide are also useful for

CC transforming any cell or organism in order to be expressed in this cell

CC or organism and result in an altered, preferably increased oil content of

CC this cell or organism. The present sequence represents the yeast

CC (Saccharomyces cerevisiae) PDAT ORF (open reading frame) nucleotide

CC sequence

XX SQ

QY Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;

Db Query Match

QY 3.0%; Score 49.4; DB 3; Length 1986;

Db Best Local Similarity 49.2%; Pred. No. 0.00044;

QY Matches 162; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

Db 428 TTGGTATGAAAGCAATGCAATGTCGCTGTCATAGCATTTGGAGATTGTCACCAACA 487  
 |||||  
 Db 815 TTGGCTATGAAACCCATTAATAATGACGAGTGTGCTGATATTTGAGGCTTGCAATTTAG 874  
 488 AATTGGAAGAGCGTGAACCTTTACTTTCACAGAGCTCAAGTTGACCTTTGAACTGCTTTAA 547  
 |||||  
 QY 875 AATTGGAAGAGCGCATAGTACTTTACAGAGCTAAAGAACAAATCGAACTGTTTCATC 934  
 548 AACTCCGTCGCGCCCTTCTATAGTATTTGCCATTCATATGGGTATATATGCTTCAGAT 607  
 |||||  
 Db 935 AATTGAGTGTGAAAAAGTTGTTTATTTGACATTTCTATGAGTCTCAATTAATCTTTT 994  
 608 ACTTCTGGAATGCTGAGGCTAGAAATTCACCA-----AAACATTAATTTGAAAGTGC 661  
 |||||  
 QY 995 ACTTATGAAATGGGTGAGGCTGAGAGGCCCTTTACGCTAATGCTGCTGCTGAG 1054  
 662 TTGATCAGCATATCCATGCTTAATTCGCTGTTGAGCTCTCTTCTTGGTCTGTTGAG 721  
 |||||  
 Db 1055 TTAAAGAAACATAGACTCATTCATTAAATGACAGAGGAGGCTTGGGCGCTCCAAAG 1114  
 722 CAATCAATCTACTCTCTCTGCTGTAACG 750  
 |||||  
 QY 1115 CAGTTCAGCTTACTTAAGTGTGAATG 1143

## RESULT 14

AAS01342

ID AAS01342 standard; DNA; 1986 BP.

AC AAS01342;

DT 31-MAY-2001 (first entry)

DE Yeast LCAT related open reading frame 1 (LRO1) gene.

DB Lactin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;

KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;

KW nutritional supplement; dairy product; food product; salad dressing;

XX yeast; Arabidopsis thaliana; LCAT related open reading frame 1; LRO1; ds.

OS Saccharomyces cerevisiae.

PN Key

FT Location/Qualifiers

FT CDS

FT 1..1986

FT /tag= a

FT /product= "LRO1"

PN WO200116308-A2.

PD 08-MAR-2001.

PF 30-AUG-2000; 2000WO-US023863.

PR 30-AUG-1999; 99US-0152493P.

PA (MONS ) MONSANTO CO.

PI Laesener M, Van Bennebaam A;

XX WPI; 2001-169010/17.

XX P-PSDB; AAU00464.

XX New isolated nucleic acid encoding plant lecithin:cholesterol  
 PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol acyltransferase  
 PT -like polypeptide, for modifying the sterol content and oil production  
 of plants.

PS Claim 5; Page 116-117; 127pp; English.

CC The present sequence encoding for yeast LCAT related open reading frame,  
 CC (LROI) is described in an invention relating to several novel  
 CC polynucleotides encoding the plant sterol acyltransferase LCAT  
 CC (lecithin:cholesterol acyltransferase-like; AAS01081-AAS01104, AAS01341)  
 CC and ACAT (acyl CoA:cholesterol acyltransferase-like; AAS01311-AAS01319).  
 CC A rat ACAT-like cDNA sequence (AAS01105) is also described. The  
 CC polynucleotides encoding LCAT or ACAT are used to produce LCAT or ACAT  
 CC polypeptides. They can also be used in a recombinant construct to  
 CC transform a host cell (preferably of a plant) or a plant. The recombinant  
 CC construct is used to increase or decrease the sterol content of the host  
 CC cell or plant. It can be used to alter oil production of the cell or  
 CC plant, preferably by increasing it. The oil of the plant or the plant  
 CC itself is used as a food product, or as nutritional or dietary  
 CC supplements, or in pharmaceutical compositions for lowering cholesterol.  
 CC The oil can be used in foods e.g. margarine, butter, cooking oil, and  
 CC dressings e.g. salad dressings, mayonnaise, cheese, processed meat,  
 CC pasta, sauces, cereals, desserts, dips, chips, baked goods, pastries,  
 CC cookies, snack bars, confections, chocolates, and beverages. The  
 CC alteration in sterol content and/or composition can also provide a plant  
 CC with tolerance to stress and insect damage

CC Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;

CC Query Match 3.0%; Score 49.4; DB 5; Length 1986;

CC Best Local Similarity 49.2%; Pred. No. 0.00044; Mismatches 161; Indels 6; Gaps 1;

DB 428 TTGGTATAGAACCAATGATTCGCTGCTCCATACGATTTGAGATTCACCAACCA 487  
 DB 815 TTGGCTATAGAACCAATGATTCGCTGCTCCATACGATTTGAGATTCACCAACCA 874  
 QY 488 AATTGGAAGCGTGACCTTACTTACCAAGCTCAAGTTGACCTTTGAACCTGCTTAA 547  
 DB 875 ATCTGAAAAGACGCATAGTACTTACCAAGCTCAAGTTGACCTTTGACCTTAA 934  
 QY 548 AACTCGGCGGCGCCCTTATAGTATTTGCCCATTCATGGGTAAATATGCTCAAGT 607  
 DB 935 AATTAGGCGTAAAAAGTTGTTTATATGACATTTAGGCGTCTCGATTAATCTTTT 994  
 QY 608 ACTTCTGGAATGCTGAGGCTAGAAATGACACCA-----AAACATTATTGAAGTGC 661  
 DB 995 ACTTTATGAATGGGTGAGGCTGAAAGGCCCTTTAGCGTAAATGATGCTGCTGAG 1054  
 QY 662 TTGATCAGATATTCATGCTTATTTGCTGTTGAGAGCTCTTCTTGGTTCTGTTAG 721  
 DB 1055 TTAAAGAACACATAGATTCATTATTAAGACAGGAGCGCTTCTGGCGCTCCAAAG 1114  
 QY 722 CAATCAATATCTCTCTCTGCTGTTAAG 750  
 DB 1115 CAGTTCAGCTCTAATTAAGTGTGAAGT 1143

RESULT 15

ADP47816 standard; DNA; 1986 BP.

ADP47816;

26-FEB-2004 (first entry)

Yeast acyltransferase nucleotide sequence SEQ ID NO:1.

XX acyltransferase; enzyme; membrane-spanning region;  
 KW active membrane independent acyltransferase; fatty acid ester;  
 KM fatty acid thioester; triacylglycerol; diacylglycerol; monoacylglycerol;

KM phospholipid; glycolipid; lysolipid; waxester; acylated carbohydrate;  
 KW acylated amino acid; structured lipid; fat-soluble molecule; yeast; gene;  
 de.

OS Saccharomyces cerevisiae.

PN WO2003100044-A1.

PD 04-DEC-2003.

PF 28-MAY-2003; 2003WO-SE000870.

PR 29-MAY-2002; 2002SE-00001581.

PR 29-MAY-2002; 2002US-0383899P.

PR 20-JAN-2003; 2003SE-00000142.

PA (SCAN-) SCANDINAVIAN BIOTECHNOLOGY RES AB.

DR Dahlqvist A, Ghosal A, Lindqvist Y, Banas A;

DR WPI; 2004-053268/05.

DR P-PSDB; ADP47818.

PT New nucleotide sequence encoding an improved acyltransferase polypeptide,  
 PT useful for producing structured lipids or fat-soluble molecules, in  
 PT removing undesirable fat or in modifying lipids in animal or plant raw  
 PT material.

PS Claim 5; SEQ ID NO 1; 91pp; English.

CC The present invention describes a nucleotide sequence (I) derived from a  
 CC nucleotide sequence encoding an acyltransferase polypeptide comprising at  
 CC least one membrane-spanning region. (I) encodes an improved active  
 CC membrane independent acyltransferase polypeptide in which at least one  
 CC amino acid residue of the membrane-spanning region has been deleted  
 CC and/or substituted as compared to the original acyltransferase  
 CC polypeptide, where the encoded active membrane independent  
 CC acyltransferase polypeptide can produce fatty acid esters and/or fatty  
 CC acid thioesters such as triacylglycerols, diacylglycerols,  
 CC monoacylglycerols, phospholipids, glycolipids, lysolipids, waxesters,  
 CC acylated carbohydrates and acylated amino acids. Also described: (1) a  
 CC nucleotide sequence molecule (II) comprising at least one promoter region  
 CC which functions in a host, where the promoter region is operably linked  
 CC to at least one (I), which is operably linked to at least one non-  
 CC translated region which functions in a host; (2) a vector comprising (II)  
 CC; (3) a host cell comprising (II) or the vector; (4) producing an  
 CC improved active membrane independent acyltransferase polypeptide,  
 CC comprising providing the above host cell and a growth medium preparing a  
 CC host cell culture, culturing the host cell culture and harvesting the  
 CC host cell culture and recovering the improved active membrane independent  
 CC acyltransferase polypeptide; (5) a polypeptide obtained by the method in  
 CC (4), where the polypeptide at least is an improved active membrane  
 CC independent acyltransferase polypeptide; (6) an oligonucleotide  
 CC specifically hybridising to (I) under stringent conditions; and (7) a kit  
 CC comprising the above polypeptide and a stabiliser. The nucleic acid  
 CC molecule (I) and the polypeptide are useful in producing structured  
 CC lipids or fat-soluble molecules, in removing undesirable fat or in  
 CC modifying lipids present in animal and plant raw material. The present  
 CC sequence is used in the exemplification of the present invention.

CC Sequence 1986 BP; 615 A; 361 C; 478 G; 532 T; 0 U; 0 Other;

CC Query Match 3.0%; Score 49.4; DB 12; Length 1986;

CC Best Local Similarity 49.2%; Pred. No. 0.00044; Mismatches 161; Indels 6; Gaps 1;

DB 428 TTGGTATAGAACCAATGATTCGCTGCTCCATACGATTTGAGATTCACCAACCA 487  
 DB 815 TTGGCTATAGAACCAATGATTCGCTGCTCCATACGATTTGAGATTCACCAACCA 874  
 QY 488 AATTGGAAGCGTGACCTTACTTACCAAGCTCAAGTTGACCTTTGAACCTGCTTAA 547  
 DB 875 ATCTGAAAAGACGCATAGTACTTACCAAGCTCAAGTTGACCTTTGACCTTAA 934

```
Oy 548 AACTCCGTCGCGCCCTTCTATAGTATTTGCCCATTCATGGGTAAATATGCTTCAGAT 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 AATTGAGTGTGAAAAAGTTTGTATATGACATTCTATGGGTCTCAGATTATCTTT 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 608 ACTTCTGGAATGCTGAGGCTAGAAATGCAACA-----AAACATTATTTGAAGTGC 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 ACTTATGTAATGGGTGAGGCTGAAGGCCCTTTACGGTAAATGATGATGCTGCTGGG 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 662 TTGATCAGCATATCCATGCTTATTTGCTGTGAGCTCCTCTTCTGCTGCTGAGG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 TTAACGAACACATAGATTCAATTATGACAGGAGGCTTCTGGGCTCCAAAGG 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 722 CAATCAAACTACTCTCTGCTGCTGTAACG 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 CAGTTCAGCTCTAATTAAGTGTGAATG 1143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 9, 2005, 04:57:01  
Job time : 936 secs

THIS PAGE BLANK (USPTO)

UD LIZI GAGGAAIAIMIBUIMACUGAUBUCILMMIUBBUICUBHUAIMWCUIGAHUCNIAVAIN

Qy 689 CTGTGAGCTCTCTTCTGTTGTTGAGGCATCAATCTCTCTGTTGTA 748  
Db 1181 ATATTAGTGTCTCTTGGTACCAAGGCTATCCAGCATTAATTTCTGGTAA 1240  
Qy 749 CG 750  
Db 1241 TG 1242

RESULT 2  
US-08-916-421B-1  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Built et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-  
Patent No. 6503729  
TITLE OF INVENTION: jannaschii  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
CURRENT FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: US 60/024,428  
PRIOR FILING DATE: 1996-08-22  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (10398)..(10398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (16385)..(16385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (19189)..(19189)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature

LOCATION: (19195)..(19195)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (23180)..(23180)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234720)..(234720)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (71684)..(71684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (85539)..(85539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)

```

Query Match      2.5%; Score 40.6; DB 4; Length 1664976;
Best Local Similarity 47.8%; Pred. No. 3;
Matches 118; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY      330 TGACAGTGTCTTTCAGCCATCACAAGATTGCATCCAGTTACATTAACAGTCCCTCTTTC 389
Db      1220217 TGGTTGTGGTTTGTGTGCTTTAGCATGTCCATTGGAGCTATTAAGAATTGATGAGTAGC 1220276

QY      390 TACTGTCTGGAAAGAGTGGCTTTAAAGTGTGTGTGAGTTTGGTATPAGAACCAATTGCAAT 449
Db      1220277 GATTAATAATGTATATGATGCTTTATTAAGAAGAGATGTTGAGATTTGGCAAGAAGATATCTCAAC 1220336

QY      450 TGTGGCTGTCCATACATGATGGAAGTTGTGACCAACCAATTTGGAAGGCGTGACCTTTA 509
Db      1220337 AGGAGCTATTAACAACCTTGAAGAAATTTAATATATTAATTAACAATAACAGTGAATTAA 1220396

QY      510 CTTTCACAAGCTCAAGTTGACCTTTGAAACTGCTTTAAACTCCGTGGCGGCCCTCTCAT 569
Db      1220397 ATTATATAGCTTACTATCTTTATGCAAAATGCAAAATTAATTCCTTAATTTTCTTAATTT 1220456

QY      570 AGTATTT 576
Db      1220457 CGTAAAT 1220463

RESULT 3
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Buit et al.

```

```

1 TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
2 Patent No. 6797466
3 TITLE OF INVENTION: jannaschii
4 FILE REFERENCE: PB2/5C1
5 CURRENT APPLICATION NUMBER: US/09/692,570
6 CURRENT FILING DATE: 2003-01-14
7 PRIOR APPLICATION NUMBER: US 60/024,428
8 PRIOR FILING DATE: 1996-08-22
9 PRIOR APPLICATION NUMBER: US 08/916,421
10 PRIOR FILING DATE: 1997-08-22
11 NUMBER OF SEQ ID NOS: 20
12 SOFTWARE: PatentIn version 3.1
13 SEQ ID NO 1
14 LENGTH: 1664976
15 TYPE: DNA
16 ORGANISM: Methanococcus jannaschii
17 FEATURE:
18 NAME/KEY: misc_feature
19 LOCATION: (28222)..(28222)
20 OTHER INFORMATION: n equals a, t, c, or g
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (28257)..(28258)
24 OTHER INFORMATION: n equals a, t, c, or g
25 FEATURE:
26 NAME/KEY: misc_feature
27 LOCATION: (84773)..(84773)
28 OTHER INFORMATION: n equals a, t, c, or g
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: (84808)..(84808)
32 OTHER INFORMATION: n equals a, t, c, or g
33 FEATURE:
34 NAME/KEY: misc_feature
35 LOCATION: (84812)..(84812)
36 OTHER INFORMATION: n equals a, t, c, or g
37 FEATURE:
38 NAME/KEY: misc_feature
39 LOCATION: (98120)..(98120)
40 OTHER INFORMATION: n equals a, t, c, or g
41 FEATURE:
42 NAME/KEY: misc_feature
43 LOCATION: (98159)..(98159)
44 OTHER INFORMATION: n equals a, t, c, or g
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: (98239)..(98239)
48 OTHER INFORMATION: n equals a, t, c, or g
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: (98266)..(98266)
52 OTHER INFORMATION: n equals a, t, c, or g
53 FEATURE:
54 NAME/KEY: misc_feature
55 LOCATION: (98343)..(98343)
56 OTHER INFORMATION: n equals a, t, c, or g
57 FEATURE:
58 NAME/KEY: misc_feature
59 LOCATION: (10398)..(10398)
60 OTHER INFORMATION: n equals a, t, c, or g
61 FEATURE:
62 NAME/KEY: misc_feature
63 LOCATION: (148948)..(148948)
64 OTHER INFORMATION: n equals a, t, c, or g
65 FEATURE:
66 NAME/KEY: misc_feature
67 LOCATION: (163385)..(163385)
68 OTHER INFORMATION: n equals a, t, c, or g
69 FEATURE:
70 NAME/KEY: misc_feature
71 LOCATION: (191989)..(191989)
72 OTHER INFORMATION: n equals a, t, c, or g
73 FEATURE:

```

```
/ NAME/KEY: misc_feature
/ LOCATION: (19195) .. (19195)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (231980) .. (231980)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234187) .. (234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234220) .. (234220)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234814) .. (234814)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309398) .. (309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309418) .. (309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312837) .. (312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312993) .. (312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (319226) .. (319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559167) .. (559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559241) .. (559241)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (600992) .. (600992)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (622708) .. (622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657081) .. (657081)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (674435) .. (674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (682442) .. (682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
```

```
/ LOCATION: (713652) .. (713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (741684) .. (741684)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (779455) .. (779455)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (855539) .. (855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (871619) .. (871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1084830) .. (1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1096846) .. (1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1119881) .. (1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1130881) .. (1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1310988) .. (1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1313224) .. (1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1349473) .. (1349473)
/ OTHER INFORMATION: n equals a, t, c, or g

Query Match      2.5%  Score 40.6; DB 4; Length 1664976;
Best Local Similarity 47.8%; Pred. No.3;
Matches 118; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY      330 TGACAGTGTCTTCACGCATTCACAGATTGATCCAGTTTACATTAACAGTCTCTTC      389
Db      1220217 TGGTTGTGTTTGTGTGCTTTAGCATGTCATTGGAGCTTAAAGATTGATGAGTAGC      1220276
QY      390 TACTGTCTGGAAGAAGTGTGCTTAAGTGTGTGAGTTTGCTATAGCAAGCAATGCAAT      449
Db      1220277 GATTAATAATGTAATGATGATTTTAAAGAGATGTTGCAAAAGAAATGATGTCAC      1220336
QY      450 TGTGCTGTTCATATCATGATTTGTCACCAACCAATTTGGAAGAGCGTGACCTTTA      509
Db      1220337 AGGAGCTATTAAACAACCTTGAAGAAATTTAAATATATTAATACAAATATACAGTAATTA      1220396
QY      510 CTTTCACAAGCTCAAGTTGACCTTTGAATGCTTTAAACTCCGTGGCGGCCCTTCAT      569
Db      1220397 ATTATATAGCTTACTATCTTTATGCAATGCAAAATTAATCTTAAATTTCTATTTT      1220456
QY      570 AGTATTT 576
      ||| |||
```



Db 1220457 CGTAATT 1220463

## RESULT 4

US-09-639-207-13/c

Sequence 13, Application US/09639207

Patent No. 6815575

GENERAL INFORMATION:

APPLICANT: Kazemi-Bafarjani, Parva

APPLICANT: Benzer, Seymour

TITLE OF INVENTION: AN ANIMAL MODEL OF POLYGLUTAMINE

TITLE OF INVENTION: TOXICITY METHODS OF USE, AND MODULATORS OF POLYGLUTAMINE

FILE REFERENCE: 06618-686001

CURRENT APPLICATION NUMBER: US/09/639,207

CURRENT FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: US 60/148,934

PRIOR FILING DATE: 1999-08-12

PRIOR APPLICATION NUMBER: US 60/148,933

PRIOR FILING DATE: 1999-08-12

PRIOR APPLICATION NUMBER: US 60/177,047

PRIOR FILING DATE: 2000-01-18

PRIOR APPLICATION NUMBER: US 60/205,720

PRIOR FILING DATE: 2000-05-19

NUMBER OF SEQ ID NOS: 69

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13

LENGTH: 486

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic DNA

US-09-639-207-13

Query Match 2.4%; Score 38.8; DB 4; Length 486;

Best Local Similarity 56.2%; Pred. No. 0.045;

Matches 73; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE: 435

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,113

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: PTZgpt-Fls

US-08-232-463-14

Query Match 2.4%; Score 38.6; DB 1; Length 7218;

Best Local Similarity 3.2%; Pred. No. 0.33;

Matches 11; Conservative 190; Mismatches 144; Indels 0; Gaps 0;

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE: 435

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,113

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: PTZgpt-Fls

US-08-232-463-14

Query Match 2.4%; Score 38.6; DB 1; Length 7218;

Best Local Similarity 3.2%; Pred. No. 0.33;

Matches 11; Conservative 190; Mismatches 144; Indels 0; Gaps 0;

NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13539  
LENGTH: 314798  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1) ... (314798)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13539

Query Match 2.4%; Score 38.6; DB 4; Length 314798;  
Best Local Similarity 46.7%; Pred. No. 4.5;  
Matches 122; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 520 CTCAGTGGACCTTTGAAACGCTTTAAACCTCGGCGGCCCTTCTATGATTTGGC 579  
DB 138072 CTACTGTAACTTTCACATCTCTTATGAAATCATTTTGTCTACCTGACCTCATTTGCT 138131  
QY 580 CATTCATGGGTAATATGCTTCTCAGATCTTCTGGAATGGCTGAGCTAGAAATTTGCA 639  
DB 138132 CATCTATGATTTGCAATATGATCTCAATCTAATTTCACTGCTAGACTAGTAAATGACA 138191  
QY 640 CCAAAACATTATTGAAGTGGCTTGAATGCAATTCATTCATTTATTTGCTGTTGAGCT 699  
DB 138192 ATGAAATAATGTTCTATTAAGTGTAGAGCTCTCTTCCATATGTAATTTTATGCTCC 138251  
QY 700 CCTCTTCTGTTCTGTTGAGGCAATCAATCTCTCTCTGTTGTAACGTTTGCCCTT 759  
DB 138252 TCTCAAGGGGCTCTTTGTTACTGAAAGTGTATTTACTTTAATGAAGCATGTGCCAT 138311  
QY 760 CCTGTTCTGAGGGAAGTCT 780  
DB 138312 GGTGAATCTTTGACATTAAT 138332

RESULT 7  
US-09-806-708B-22/c  
Sequence 22, Application US/09806708B  
Patent No. 6784342  
GENERAL INFORMATION:  
APPLICANT: The University of British Columbia  
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants  
FILE REFERENCE: 4810-58741  
CURRENT APPLICATION NUMBER: US/09/806,708B  
CURRENT FILING DATE: 2001-04-03  
PRIOR FILING DATE: 1999-08-04  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patent version 3.0  
SEQ ID NO 22  
LENGTH: 1141  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (1) ... (1141)  
OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. FAEI promoters  
US-09-806-708B-22

Query Match 2.3%; Score 38.2; DB 4; Length 1141;  
Best Local Similarity 9.6%; Pred. No. 0.13;  
Matches 73; Conservative 280; Mismatches 405; Indels 3; Gaps 1;

QY 488 AATGGAGAGCGTACCTTTACTTTCACAACTCAAGTGCCTTTGAAATGCTTTAA 547  
DB 884 RHHTTCRTTKNNNNNNNATVYWHHAARRWMAAMTETNNNNNNNNNNACRNTRTWMA 825  
QY 548 AACTCCGTTGGGCGCTTCTATAGTATTTGCCATTCATGGTATATAATGCTTCAGAT 607  
DB 824 BKKSKWNN 765

QY 608 ACTTCTGGAATGGCTGAGCTAGAAATTTGACCAACCAAAATTAATTTGAAGTGGCTGATC 667  
DB 764 THTDWCYTMNNNTWMDWMTTMBTTTRNNMTTSTMTNNNNNNNNNNNNNNNNNNNNNNNNNNNN 705  
QY 668 AGCATATCCATGCTTATTTCCGCTGTGGAGCTCCCTCTTGGTTCTGTTGAGGCAATCA 727  
DB 704 AATATNNMGCMNNNDARNTNNYTVWRBRMNTNTKRWYSTTRRHYYTATATNNNNNNNNNN 645  
QY 728 AATCTACTCTCTGGTGAAG--TTGGCCTTCCGTTCTTGAGGAAGTGGCTGGT 784  
DB 644 NNNNNNSCCTCTTATTTTMMTKGDMTRKVKYKRDITCTTVDWADSWVMYANMRCR 585  
QY 785 TGTGTCCAAATCTTTTTCGCTGATATGCTTATGCTTATGCCATTTTCAAGAATTTGCAAG 844  
DB 584 DVTYRNNYTKSYAHSYWYSSNNAMWYRXYRARSNNWSMAMTTTRNNMMGGBVYMRWAG 525  
QY 845 GTGATAACAATCTCTGAGCGCATTTTGTGGGGGTGCTCAAAGAAGATAAGCCGGAT 904  
DB 524 TMMRHNNNNNTDTRYYWMMKRWABTTTVDSCNKAWSMRGNMWRAMKRWMAANDAG 465  
QY 905 ACCACTGTGATGAAGAGATATCATCAATAATTTCTGGCTGGCCGCAATAATATTA 964  
DB 464 AMDHTTYMGMNTMMKRAAMMMAMKCRALYCCNNNNNNRACVWHKHMRTKTKYMKKA 405  
QY 965 ACATTGAAATCTTTCACCTAGCTGACGTTACAGAAAGCTCTAGTCAACATGACCAAGATG 1024  
DB 404 ACNNNNBKAATMRVAMMYSRDNTTIDWMMTSDMBHMYTVDTYMMRAMNNNNNNNNNN 345  
QY 1025 AATGCGCTTCCACCTTTTGTCTTTCACAGCCCGGAACTAGCAATGAGACTCTT 1084  
DB 344 CTTSSMMMMHMHNTCTYGNNTGSAVBMAISMAAGASBNVYNNCMTTYYGKTMT 285  
QY 1085 TCAAGCAATGAAGACTATGACCCAGATAGCAAGAGATGTATCAACAGTTAAAGAGT 1144  
DB 284 NNNNNNNKMYRTKTVAMCNRNRYDDTAVTWBKNTKYCYAVBYWYBMHMKHHBWR 225  
QY 1145 TGTATCATGATGACCTGTTTATCTCTGACTCTCTGGAGAGACCACTATATAA 1204  
DB 224 ABHRSMMWVVCNRKYVWSHMYHAMRYKBAVAVGNNNNKDDRAHHHMCATNNNNMMW 165  
QY 1205 ATGTATTTGATATATGCTGCTCATCTTAAGACAGAGTT 1245  
DB 164 WYAVHHHKKKGKAAWTKNKAABRDDBAHVKTWYMYRDY 124

RESULT 8  
US-09-949-016-2027  
Sequence 2027, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03/231,498  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2027  
LENGTH: 4465  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-2027

Query Match 2.3%; Score 37.6; DB 4; Length 4465;  
Best Local Similarity 57.8%; Pred. No. 0.52;  
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```

Oy      998  CAGCCTCTGTCAAACATGACCAAGCATGGAATGGCGCTTCCACCCCTTTGTCTTTACAG 1057
Db      1216  CAGAGATTCTGAATATCAAGCATCTGCAGAAAGATGGGACCCCATATATGCTTTCCAC 1275
Oy      1058  CCCGTGAACCTAGCAGATGGAGCTCTTTTCAAGCATAGAAAGACTATGACCCAGAT 1113
Db      1276  ACTTTTACCAAGCATGATGAGAGGTTTGTTCTGCGCATGAAAGGATGACCCCAAT 1331

RESULT 9
US-09-621-976-2439/c
; Sequence 2439, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2439
;
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 146..310
; US-09-621-976-2439

```

Query Match	2.2%	Score 36.6	DB 4	Length 581
Best Local Similarity	47.9%	Pred. No. 0.28		
Matches	105	Conservative	0	Mismatches 114; Indels 0; Gaps 0;
QY	596	ATGTCCTCAGACTCTTCTGGAATGCGTAGGCTAGAAATTGCACCAAAACATTATTGGA	655	
Db	395	AAGCTTCAGGTCCTCCTCTATAGAGAAAGTGCAGTTCAAGCCACATCAGAGATTAA	336	
QY	656	AGTAGCTGATCAGCATATCCATGCTTATTTCGCTGTGAGATCCCTCTTCTGGTCTG	715	
Db	335	TATAGCGTGAAATCTCTTTCCATTATTAATTCGCCTTAGTAATCTTTTGTGATTTTG	276	
QY	716	TTGAGGCATCAATCTACTCTCTCTGATGTAACGTTTGCCCTTCTGTTCTAGGAA	775	
Db	275	TATTGGAATTAGAGGTTGCGCAATTATGAAGTCAGATGTCAGTGACAGTGTCAAGCGT	216	
QY	776	CTGCTCGGTGTGTGCAAATCTTTGGCGTGCATGCT	814	
Db	215	ATCCCAATTGTGCTGTGTTCTCTGTCTCTCTTGT	177	

```

RESULT 10
US-09-621-976-2813
: Sequence 2813 Application US/09621976
: Patent NO. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSET 054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 2813
: LENGTH: 832
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

```

Query Match	2.2%;	Score 35.6;	DB 4;	Length 832;
Best Local Similarity	11.5%;	Pred. NO. 0.76;		
Matches	37;	Conservative 150;	Mismatches 134;	Indels 1;
			Gaps	1;

Oy		363	TCCAGGTATCAATAACAGGCTCCTTTCTACTGTCGTGAAGAAGGGCTTAAGTGCGTGT	422
Db		26	MMYKMYMKTMYMRBRKKKAMMKMYMTWMTWMTAYAMGTYYKKMACRTTKKKKKKG	85
Oy		423	TGAGTTTGGTATAGAACAATGCATTGTGCGTGTTTCATACGATTGGAAATTGTCAAC	482
Db		86	YMMMYMGWRSSYNAMMTWTWTGYAAYRSSMYMRIRWKKKAKYRKTTYSSKKMTWMK	145
Oy		483	AACCAAATTTGAAGAGCGGTGACTTTTACTTTACAAAGCTCAAGTTGACTTTGAAACTGC	542
Db		146	RW-KKAATWTMMKXTTYWAATRRVMMMCMTBMRBASMWYICMMWGAKRMSTIMRSRYAS	204
Oy		543	TTTTAAAATCCGTGCGGCCCTTCTTATAGTATTTGGCCATTCAATGGTAAATATGCTT	602
Db		205	ARSARKRCYCSGSWGMASMKYMMRMMRWMGWVTGAGMAWRASCMMRRRYAKSKTSYSVM	264
Oy		603	CAGATTACTTTCGATAGTAGCTGAGGCTAGAAATTTGCAACAAAGATTTATTTGAAGTGCT	662
Db		265	IMCMTRSMKYCYTAAARMTGYCYCRKGGMMGRBGRHWYASKYTMWKRMMCMABMYRYSTG	324
Oy		663	TGATCAGCATATCCATGCTTAT	684
Db		325	RASMNRMRWYTMMKMKWKIAM	346

```

1 RESULT 11
2 US-08-809-254A-4
3 : Sequence 4, Application US/08809254A-4
4 : Patent No 6660852
5 : GENERAL INFORMATION:
6 : APPLICANT: KSHI ET AL
7 : TITLE OF INVENTION: PROBE FOR DIAGN
8 : FILE REFERENCE: 19036/33767
9 : CURRENT APPLICATION NUMBER: US/08/8
10 : PRIOR FILING DATE: 1997-05-16
11 : PRIOR APPLICATION NUMBER: PCT/JP95/
12 : PRIOR FILING DATE: 1995-10-02
13 : PRIOR APPLICATION NUMBER: JP 256348
14 : PRIOR FILING DATE: 1994-09-30
15 : NUMBER OF SEQ ID NOS: 9
16 : SOFTWARE: Seqin version 3.1
17 : SEQ ID NO 4
18 : LENGTH: 5829
19 : TYPE: DNA
20 : ORGANISM: Artificial sequence
21 : FEATURE:
22 : OTHER INFORMATION: Synthetic probe
23 US-08-809-254A-4

```

Query Match	2.1%	Score 35;	DB 4;	Length 5829;
Best Local Similarity	49.2%;	Pred. NO. 4.6;		
Matches	92;	Conservative	0;	Mismatches 95;
			Indels	0;
			Gaps	0;

Qy	642	AAAAATTATTTGAAGTGCTTGATACGATATCCGCTTAATTCGCTGTGGAGCTCC	701
Db	3443	AAAGCATCTAATTAAAGCTTTTCAAGACTAATAATATGGCGTGTTCGATTTGAAAGAC	3502
Qy	702	TCTTCTTGTCCTCTGTGAGCAATCAATCTCTCTGCTGTAAAGTTGGCTTCC	761
Db	3503	ATTTCATTAATTATTACTGTAAACCAACATCTGGCTTACTGTGAATTCGCTTGTCC	3562
Qy	762	TGTTTCTGAGGAACCTGCTCGGTGTGTCTCAATTCCTTTGGCGTGCATTTGTGCTTAT	821
Db	3563	TGCCAGGTTCCTCGTGTGTTTGAGTAGTACACAGCTGTCTTCGTCGCTGACTGGATVAA	3622

QY 822 GCCATT 828  
| | | | |  
Db 3623 GCCATT 3623

## RESULT 12

US-08-961-527-41/c  
; Sequence 41, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunach  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9828 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-41

Query Match 2.1%; Score 35; DB 3; Length 9828;  
Best Local Similarity 49.2%; Pred. No. 6.6;  
Matches 92; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 642 AAAACATTATTGAAGTGGCTGATCAGCATATTCATGCTTATTCGCTGTGAGCTCC 701  
| | | | |  
Db 4114 AACGATACATATTAAGCTTTTCAGACCTAATAATATGCCGTGTTGATTTGAAAGAC 4055  
| | | | |  
QY 702 TCTTCTTGATTCGTGAGGCAATCAATCTCTCTGTGATGAAGCTTTGGCTTCC 761  
| | | | |  
Db 4054 ATTCATATATATTTTACTGTATCAAGCATCTGGCTACTGTGAATTTGGCTTGC 3995  
| | | | |  
QY 762 TGTTCGAGGAACTGCTCGGTGTTGTGCAATTTCTTTGGCTGTCATTGTGGCTTAT 821  
| | | | |  
Db 3994 TGCCAGTGTTCGCTGCTGTTGAGGTAGTACAGCTGTTCCGTCGCTGACGTGATMAA 3935  
| | | | |  
QY 822 GCCATT 828  
| | | | |  
Db 3934 GCCATT 3928

RESULT 13  
US-09-949-016-17057  
; Sequence 17057, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,766  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17057  
; LENGTH: 154023  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17057

Query Match 2.1%; Score 35; DB 4; Length 154023;  
Best Local Similarity 51.8%; Pred. No. 44;  
Matches 103; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 908 ACTGTGATGAAGAGATATCAATCAATAATTCGTGCTGGCCGCAATAATATTATTAACA 967  
| | | | |  
Db 147507 ACAGTATTCACAGATATATATATCAACATCTGCTGTGT -GCCAAGCAGTGTATTATCA 147565  
| | | | |  
QY 968 TTGAATTCCTTCCACTAGCGCTTACAGAAACAGCTCTAGTCAACATGACGATGGAAT 1027  
| | | | |  
Db 147566 CTTCATTATTAACCTGTGAAGAAATGAATGTATACATCAAGTGTGGAG 147625  
| | | | |  
QY 1028 GTGGCTCCCAACCTTTGTCTTTACAGCCCGTGAAGTCAAGATGAGACTCTTTTCA 1087  
| | | | |  
Db 147626 GAAGCATATCCATGTATATTTTGTCTGGAATATGATTAATTTCAAAAGTTGCCCTTCCA 147685  
| | | | |  
QY 1088 AAGCAATGAAGACTATGA 1106  
| | | | |  
Db 147686 AACAAAGTTGATATATTA 147704

RESULT 14  
US-09-710-279-2485  
; Sequence 2485, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 2485  
; LENGTH: 462  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2485

Query Match 2.1%; Score 34.8; DB 4; Length 462;  
Best Local Similarity 50.6%; Pred. No. 0.95;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 865 CATTTTCTGGGGGCTGCAAGAAAGATAGAGCGGCTATACCACTGTGATGAAGAGGAA 924  
| | | | |  
Db 148 CATATTTCGCCCTCCCTGCTGAGAGCAATATCAATATGATTAATTAATCTTACGCAACA 207  
| | | | |  
QY 925 TATCAATCAAAATATTTCTGGCTGGCCGCAAAATATATTAATTAACATTTGCACT 984

DB 208 TCTCAATTAGACTCAAGAGTAAGACTAAATGATTATATAGAAACAGAAATTTATTAACACT 267

QY 985 AGCGTTACGAAAGAGCTCTAGTCAACATGACCGACATGGAATGTG 1030

DB 268 ATGATTTAAAGATCGTATATACTAATAAAGACACACGAGATTGTG 313

## RESULT 15

US-09-134-001C-1521  
; Sequence 1521, Application US/09134001C

; Patent No. 6380370  
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucet-Re-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 1521  
; LENGTH: 477

; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1521

Query Match 2.1%; Score 34.8; DB 3; Length 477;  
Best Local Similarity 50.6%; Pred No. 0.97; Mismatches 82; Indels 0; Gaps 0;

Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 865 CATTTTCTGGGGGTCTGCAAGAAAGATAGCGGTATACCACTGTGATGAAGAGGA 924

DB 163 CATATTATCGCTGCTGTGAAGGCAATCAATATGATATCAATTACTTAGCACA 222

QY 925 TATCAATCAAAATATTCTGGCTGGCCGACAAATATTATTACATTGAAATTCCTTCCACT 984

DB 223 TCTCAATTAGACTCAAGAGTAAGACTAAATGATTATATAGAAACAGAAATTTATTAACACT 282

QY 985 AGCGTTACGAAAGAGCTCTAGTCAACATGACCGACATGGAATGTG 1030

DB 283 ATGATTTAAAGATCGTATATACTAATAAAGACACACGAGATTGTG 328

Search completed: November 9, 2005, 01:39:27  
Job time : 313 secs

THIS PAGE BLANK (user)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 9, 2005, 01:05:48 ; Search time 1389 Seconds  
(without alignments)  
9770.175 Million cell updates/sec

Title: US-09-651-651-4

Perfect score: 1641

Sequence: 1 atgggagcgcgaatcgaatc.....ttgataaagcgggtatctaa 1641

## Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database : Published Applications NA:\*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
27: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
28: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	591.8	36.1	1719	20	US-10-437-963-76086 Sequence 76086, A
2	252.6	15.4	1760	19	US-10-424-599-16008 Sequence 16008, A
3	215.8	13.2	528	19	US-10-424-599-134518 Sequence 134518, A
4	193	11.8	924	21	US-10-425-115-137717 Sequence 137717, A
5	187.8	11.4	848	21	US-10-425-115-93907 Sequence 93907, A

6	173.8	10.6	1383	21	US-10-425-115-93909 Sequence 93909, A
7	120.8	7.4	1433	19	US-10-424-599-17331 Sequence 17331, A
8	72.2	4.4	380	19	US-10-424-599-57573 Sequence 57573, A
9	68	4.1	2326	22	US-10-883-760-45 Sequence 45, Appl
10	56.6	3.4	369	21	US-10-425-115-57116 Sequence 57116, A
11	51	3.1	962	21	US-10-425-115-61794 Sequence 61794, A
12	44	2.7	374	21	US-10-425-115-93908 Sequence 93908, A
13	38.8	2.4	486	18	US-10-465-217-13 Sequence 13, Appl
14	38.6	2.4	325446	21	US-10-719-993-6824 Sequence 6824, Ap
15	38	2.3	1545	19	US-10-424-599-56537 Sequence 56537, A
16	37.6	2.3	394	18	US-10-242-559A-23539 Sequence 23539, A
17	37.6	2.3	394	19	US-10-085-783A-23539 Sequence 23539, A
18	37.6	2.3	473	15	US-10-060-097-143-226 Sequence 226, App
19	37.6	2.3	1292	24	US-10-773-226-198 Sequence 198, App
20	37.6	2.3	1292	24	US-10-921-235-35 Sequence 35, Appl
21	37.6	2.3	2329	24	US-10-794-899-107 Sequence 107, App
22	37.6	2.3	3737	17	US-10-240-965-140 Sequence 140, App
23	37.6	2.3	11477	26	US-11-097-143-24160 Sequence 24160, A
24	37.6	2.3	12600	26	US-11-097-143-24163 Sequence 24163, A
25	37.6	2.3	12600	26	US-11-097-143-24169 Sequence 24169, A
26	37.6	2.3	12600	26	US-11-097-143-24163 Sequence 24163, A
27	36.6	2.2	1939	21	US-10-425-115-51256 Sequence 51256, A
28	36.4	2.2	296	21	US-10-357-930-18841 Sequence 18841, A
29	36.2	2.2	3325	20	US-10-437-963-65323 Sequence 65323, A
30	36	2.2	328	22	US-10-275-323A-78 Sequence 78, Appl
31	36	2.2	735	20	US-10-437-963-38267 Sequence 38267, A
32	36	2.2	2395	20	US-10-437-963-38278 Sequence 38278, A
33	36	2.2	2433	18	US-10-321-802-21 Sequence 21, Appl
34	35.8	2.2	654	13	US-09-925-065A-718981 Sequence 718981, A
35	35.8	2.2	654	13	US-09-925-065A-718982 Sequence 718982, A
36	35.8	2.2	654	13	US-09-925-065A-718983 Sequence 718983, A
37	35.8	2.2	654	13	US-09-925-065A-718984 Sequence 718984, A
38	35.8	2.2	266145	22	US-10-856-118-1 Sequence 1, Appl1
39	35.6	2.2	568	13	US-09-925-065A-114201 Sequence 114201, A
40	35.6	2.2	576	13	US-09-925-065A-665799 Sequence 665799, A
41	35.6	2.2	2356	9	US-09-822-830A-324 Sequence 324, App
42	35.6	2.2	4089	9	US-09-764-877-3884 Sequence 3884, Ap
43	35.6	2.2	4089	18	US-10-242-515-1884 Sequence 3884, Ap
44	35.6	2.2	17703	18	US-10-257-166-34 Sequence 34, Appl
45	35.6	2.2	43680	17	US-10-257-378-3 Sequence 3, Appl1

## ALIGNMENTS

```
RESULT 1
US-10-437-963-76086
; Sequence 76086, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 76086
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76111C.1
US-10-437-963-76086
Query Match      36.1%; Score 591.8; DB 20; Length 1719;
```

Best Local Similarity 65.2%; Pred. No. 1.9e-173;  
Matches 1029; Conservative 0; Mismatches 382; Indels 168; Gaps 4;

```

OY 108 CGACTACTCGAAGCTATCGGATATATCATTCGCGGATTTGGGTGAGCGGACTACGAGC 167
DB 123 CGACTACCGGAGGCTCTCGGGATATATATCCCGGGTTCGGGTGACGACGACTGGGCGC 182
OY 168 GTGGTCACTCTGATCTGTCCATACATCTCCGTGGAATTCAATCCGCTCGACCTGTATG 227
DB 183 GTGGTCTGTGGATTTGGCCCTTACTCGCTTCGACTTCAACCCCGCTGATCTCGGTG 242
OY 228 GCTAACAACCACTAAGCTTTCTTCTGCTGTCACTGCTGTGTTAAGTATGCTGTAGA 287
DB 243 GCTCAACACCGCAAGCTCTTATCTGCTGTCAATGCTGCTCAATGATGATCTCTCGA 302
OY 288 TCCTTATATCAACAGACATCCCGAGTGAAGTACGCGGCTGACAGAGTCTTTACG 347
DB 303 CCTTACACACGAGCGATCATCCAGAGTCAAGGCTTGACAGTGTCTTTCTTCTG 362
OY 348 CATCAACAGATTGATCCAGGTTACATTAACAGTCTCTTTCTACTGTCTGAGAAAGTG 407
DB 363 AATTACAGAGTTAGACCTGTTATATATATCTGCTCTTTCTTCAAGTGTGAGAAAGATG 422
OY 408 GCTTAAGTGTGTGTGAGTTGTGATAGAACCAATGCAATGTGCTGTCTCATACGA 467
DB 423 GGTCAAAATG----- 431
OY 468 TTGAGATTGTCAACCAACCAATTGGAAGCGGTGACCTTATCTTACAGAGCTCAAGT 527
DB 432 -----GTT 434
OY 528 GACCTTTGAAACCTGCTTTAAACCTCGGTGCGGCGCTTCTATAGTATTTGCCATCAAT 587
DB 435 AACTTTGAAACCTGCTTTGAAACCTCGAGGAGGCGCTTCTTATAGTGTGCTCATTCAT 494
OY 588 GGGTAAATATGCTTCAATATCTTTCTGGAATGCGTGAAGCTAAGAAATGACCAAAAGA 647
DB 495 GGGTAAATATGCTTCAATATCTTTCTGGAATGCTTGAAGCTAAGAAATGCTCCCAAGA 554
OY 648 TTAATTTGAAGTGGCTTGTATCAGCATATCCATCTTATTTGCGTGTGAGCTCTCTTCT 707
DB 555 TTACATCCGATGGCTTTGACGAACTATATCATCTTGTGAGTGTGAGCACTCTTCT 614
OY 708 TGGTCTGTGAGCAATCAATCTCTCTGTGTGTAAGCTTTGCTGTCTTCTGTTTC 767
DB 615 TGGTCTGTGAGCAATCAATCTCTCTGTGTGTAAGCTTTGCTGTCTTCTGAGTGA 674
OY 768 TGAGGGAATGCTGCTGCTGTGTGCTCAATCTTTTGGCGTGTGATTTGGCTTATGCCANT 827
DB 675 CGAGGGAACAGACGATTTGATTTAAATGCAATTTGGTTCATCTTATGCTCATGCCCTT 734
OY 828 TTCAAGAATTGCAAGGCTGATACACATCTGAGCGCATTTTCTGGGGTGTGCAAA 887
DB 735 CTCAGAAATTTGCAAGCTGATATATATCTGAGAGCATTTCTTTGAGGGAAGGAA-- 792
OY 888 GAAGAATAAGCGCGTATACCACTGTGATGAAGAGAAATATCAATAAATATTTGCGCTG 947
DB 793 -GGTTCACACACAGACGATGTGATGAATGAATGAATGAATGAATGAATGAATGAATGA 851
OY 948 GCCGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1007
DB 852 GCCGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 911
OY 1008 CAACATG-----ACGACATGAGATTTGGCTTTCCACCCCTTTT 1046
DB 912 CATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 971
OY 1047 GTCTTTCAAGCCGCTGAATGAGAGTGAAGTCTTTTCAAGAAATGAAGATATGA 1106
DB 972 GTCAATTTTCTGCTAGAGGAGTTCAGATGTGCTCTGTTTAAACCATTAAGATTTGGA 1031
OY 1107 CCCAGATGAAGAGATTTTACACAGTTAAAGAGTTGTATCANTGATGACCTGTTT 1166

```

```

DB 1032 CCTTCAGACATAGGCGCTTATTCATCAGCTTGAAGAGTACTACAGGGTGTATCCAGTTCT 1091
OY 1167 TAATCCTCTGACTCCTTGGGAGAGACGACCTATTAATAATGATATTTGCTATATAGTGC 1226
DB 1092 TAATCCTCTGACCCCTGGAGAGAGCCCCCAATTAAGATGTGTTCTGCTATATATAGGCT 1151
OY 1227 TCATTAAGAAGAGGTTGGTTATTAATTTTGGCCCAAGTGGCAAACTTTATCTGATTA 1286
DB 1152 TGATTCMAAGACTGAGTATGATTTATTTTTCACCGAGTGGAAAAACATATCCAGATTA 1211
OY 1287 TTGATTCATCAGGATATCATTTTACGAAACCTGAAGGTTCCCTCGTGTCAAG----- 1338
DB 1212 CTGATTAATTAATCTGATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1271
OY 1339 -----TGTGAACCTGTGTTGATTTGGAACGCTGAGC 1370
DB 1272 AATAGTACACAAATTTATTTCTTTCAGATTCAGATTTCTGTACTGAGAAAGCCCAACAA 1331
OY 1371 TATTAATGAGGAGATGAGACGATACCTTATCATTTCACTCTTGTGTGCAAGAAATTTGGCTGG 1430
DB 1332 TTCCAGCGGAGACGGAACGATATCTTCAATTTCTCTCATGTGTGCAAGAACTGGCTTGG 1391
OY 1431 ACCTTAAGTTAAACATTAACAAATGCGTCCGACGCAAGAACAGATGGAACGACATCATGT 1490
DB 1392 GCCAAAGTGAACATTAATCTAGGCGCTCCGACGCAAGAACATATGATGTGATTAACAAC 1451
OY 1491 GGAATTAATGTTGATCATGAGATGAGTGGTCAAGCATCATTAATCAATGACAAAGAACACC 1550
DB 1452 AAAGTGAATGTGACACCAACATGATGAGGCAAGCATATCCCAACATGAGAGAGCTCC 1511
OY 1551 AAGGTTAAGTACATTAACCTTTTATGAAGACTCTGAGAGCATTTCCGGGAAAGAACCCG 1610
DB 1512 ACATGTGAAGTACATTAACCTTTTGAAGATGCTGAGAGCATTTCCGGGATGAGAACAGC 1571
OY 1611 AGCTGGGAGCTGATTA 1629
DB 1572 GGTCTGGAGCTGTATTA 1590

```

```

RESULT 2
US-10-424-599-16008
; Sequence 16008, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 16008
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_114461C.1
US-10-424-599-16008

```

```

Query Match 15.4%; Score 252.6; DB 19; Length 1760;
Best Local Similarity 76.7%; Pred. No. 2.9e-67;
Matches 309; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
OY 1227 TCATTAAGAAGAGGTTGTTATTAATTTTGGCCCAAGTGGCAAACTTATCTGATTA 1286
DB 76 TGACTTTAAAAAATGATGTTGTTACTATCTATGCTTCAAGCGCAAGCTTACCTGATTA 135
OY 1287 TTGATCATCAAGGATATCATTTTACGAAACCTGAAGGTTCCCTGCTCAAGGCTGGAAC 1346
DB 136 CTGATCATTAAGGATGTGTTATGAGTTTGAAGGATCTTAAATCTCAAGGCTCAGGGA 195

```



```
QY 1347 TGTGTTGATGGAGAGCTGAGACCTTAATCTGGGAGTGAAGCGTACCTTATTCCTACT 1406
DB 196 TCTGTGTGAAGGAGACCTTGAGACCAATTAAGTGCATGAGACCGTCCATATCTCCCT 255
QY 1407 CTCTTGTCAGAAATTTGGCTCGGACCTTAAGTTAACAATGAGCTCCCGACGAGA 1466
DB 256 TTCCTGTGCAGAAAACCTGGCTTGAGACGAAAGTGAACATTAACAAAAGCTCCACAGTCAGA 315
QY 1467 ACAAGATGAGAGGACGTACATGTGAACTAAATGTTATCATGAGCATGGTCAAGAT 1526
DB 316 GCATGATGCTCAGATGTCAATTAATTAATGATGGAACATCACCATGAAGAAGATAT 375
QY 1527 CATAGTAAATGACAAAGGACCAAGGGTTAGTACATTAACCTTTTGAAGACTCTGA 1586
DB 376 CGTTCCAAACATGACAAATACCAAGGGTGAAGTATATTAACATTAATTAAGATTCGA 435
QY 1587 GAGCATTTCCGGGAGAGAGAACCGCAGTCTGGAGCTTGATTA 1629
DB 436 AAGTCTTCTGGAAAGAGACAGCAGTCTGGAGCTTGATTA 478
```

## RESULT 3

```
US-10-424-599-134518
; Sequence 134518, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; NUMBER OF SEQ ID NOS: 265684
; SEO ID NO 134518
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9247C.1
US-10-424-599-134518
```

```
Query Match 13.2%; Score 215.8; DB 19; Length 528;
Best Local Similarity 80.7%; Pred. No. 4.4e-56;
Matches 264; Conservative 0; Mismatches 62; Indels 1; Gaps 1;
```

```
QY 108 CGACTAATGGAAGCTATCGGATTAATCAATCCGGATTTGGCTGACGACGAGTACAGC 167
DB 202 CCACTACTGAAGCTCTCCGGCATTAATCCCGGCTTGGCTCAGCTCAGCTCCGAC 261
QY 168 GTGTGTCATCTTGAAGTGCATACACTCGTGTGACTCAATCCGCTCGACTCGTATG 227
DB 262 ATGTCATCTTCTGATCTCTTACTCTCCGCTGATTTCAACCTCTGACTTGTGTCG 321
QY 228 GCTAGACCACTAAGCTTTCTTCTGCTGTCAACTGCTGTTAAAGTGTATGCTGTAGA 287
DB 322 GCTGCACACCAACCACTTCTTCTGCTGTCAATGCTGCTTAAGTGAATGATGATGGA 381
QY 288 TCCCTATATCAAAAGACATCCGAGTGAAGTGAACGCGCCGAGAGAGTCTTTCACG 347
DB 382 TCCCTTAACAACGACAGATCTCTGATGAAGTCCCGTCGGAATGAGTGTCTTCTG 441
QY 348 CATCAAGATTTGATTCAGAGTTACATTAACAGGTCTCTTTC-TACTGTCTGGAAGAGT 406
DB 442 CATTAAGAACTTGAATCCAGGTTATATTAACAAGACCTCTTCACTTCAAGTTGGAAGAT 501
QY 407 GCGTTAAGTGTGTGTGATTTGTTA 433
DB 502 GGAATTAAGTGTGTATTAATTTGGCA 528
```

## RESULT 4

```
US-10-425-115-137717
; Sequence 137717, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425, 115
; NUMBER OF SEQ ID NOS: 369326
; SEO ID NO 137717
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_57075C.1
US-10-425-115-137717
```

```
Query Match 11.8%; Score 193; DB 21; Length 924;
Best Local Similarity 79.2%; Pred. No. 8.6e-49;
Matches 229; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```
QY 239 CTAAGCTCTTCTGCTGTCACTGCTGTTAAGTATGAGTGTAGATCTTATATC 298
DB 590 CTCGTGTTTCTCTGCGGTAATTTGCTGCTTAATGATGCTGTGAACCTTAATC 649
QY 299 AACAGACCATCCGAGTGTAAAGTACAGGCTGACAGTGTCTTTCAGCATCAGAA 358
DB 650 ACATGACCATCCGAAATGCAAGTCAAGCTGATGAGTGTCTTCTGCAATTAAGAGC 709
QY 359 TGAATCCAGGTATCAACAGGCTCTTCTTACTGTCTGAAAGAGTGGCTTAAGTGT 418
DB 710 TGACCTGTGTATTAACAGGCTCTCTCTTCTTCAATGATGAAGATGGTCAATGT 769
QY 419 GTGTGAGTTTGTATGAAGCAAAATGCAATTTGCTGTTTCAATGATGAGATTTGT 478
DB 770 GTGTGAGTTTGTGATGAAGTGAATGCAATTTATCGTGTTCGATATGATGAGACTGC 829
QY 479 CACCAACCAATTTGAAGCGTGAACCTTTACTTTCAACAGCTCAAGT 527
DB 830 CCCCATCAATGCTTGAAGAGAGATCTGTACTTTCACAAAATTAAGTT 878
```

## RESULT 5

```
US-10-425-115-93907
; Sequence 93907, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425, 115
; NUMBER OF SEQ ID NOS: 369326
; SEO ID NO 93907
; LENGTH: 848
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_185630C.1
US-10-425-115-93907
```

```
Query Match 11.4%; Score 187.8; DB 21; Length 848;
```



Query Match 4.4%; Score 72.2; DB 19; Length 380;  
Best Local Similarity 61.2%; Pred. No. 3.2e-11;  
Matches 167; Conservative 0; Mismatches 103; Indels 3; Gaps 3;

1234 AAGACAGAGTTGTTATTTACTTTGGCCCAAGTGGCAACCTTATCTGATTAATGGATC 1293  
109 AAGAAATAGAGCGGTACTACTATGCTCTTAGCCCGTAGACCTTACCTTGAACATGGATC 168  
1294 ATCAGGATATCAT-TTACGAAACTGAAAGTTCCCTGCTGTCAAGGCTGGAACTGTGTGT 1392  
169 ATTACTGTTGGGATACAGATAGCCAGATGCTCTTGATCAGAGCAAGAAATCCGCT 228  
1353 TGATGGAGCGTGGACCTATATACCTGGGATGAGAGGTACCTATCATCTCTCTTG 1412  
229 TGAAGGAAACCTGTATCCATAGC-CGGCATGAGAGCGCCCACTTATCCCTTTTCTG 287  
1413 GTGCAAGATTGGCTCGACCTTAAAGTTAACATA-ACAAATGGCTCCCGCAGCAGAACAG 1471  
288 GAGCAAGAAACAGGCTTGTGTCGGATGTGACATATACAGAGCCCGTGTGAGCAGG 347  
1472 ATGGAAGCAGCTACATGTGAACTAATGTTG 1504  
348 AAGGTCAGATGTTCACCTTAATGGATGTGG 380

## RESULT 9

US-10-883-760-45  
; Sequence 45, Application US/10883760  
; Publication No. US20050035174A1  
; GENERAL INFORMATION:  
; APPLICANT: Vadav, Narendra  
; APPLICANT: Zhang, Hongxiang  
; TITLE OF INVENTION: ACYLTRANSFERASES FOR ALTERATION OF POLYUNSATURATED FATTY ACIDS  
; TITLE OF INVENTION: AND OIL CONTENT IN OLEAGINOUS YEASTS  
; FILE REFERENCE: CL2302 US NA  
; CURRENT APPLICATION NUMBER: US/10/883,760  
; CURRENT FILING DATE: 2004-07-06  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 45  
; LENGTH: 2326  
; TYPE: DNA  
; ORGANISM: Variowia lipolytica  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (2271)..(2271)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-883-760-45

Query Match 4.1%; Score 68; DB 22; Length 2326;  
Best Local Similarity 49.4%; Pred. No. 2.1e-09;  
Matches 176; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

395 TCTGAAAAGTGGCTTAAGTGTGTGTTGATTGTAAGCAATGCAATTGTG 454  
1016 TGTGGAACAAGCTGCTCGAAGCTGCTGTATTGGAATACATACGATACATGCTG 1075  
455 CTGTTCCATAGATTGAGATTGTCCCAACCAATTGAAGAGCTGACCTTTACTTTC 514  
1076 CTGCGCGGTACACGTGAGACGTCTTACCCTGATTTGGAGACCGAGACGATATCTTC 1135  
515 ACAAGTCAAGTTGACCTTTGAACCTTTAAACCTCCGTGCGCGCCCTTCTATAGTAT 574  
1136 CCAAGTCAAGCTTAAATGAAAGAGACTAAGCGTATGACAGGTGAAGACAGTTCTTA 1195  
575 TTGCCATTCATAGGTAATTAATGCTTCAGATTAATTTGGAATGGCTGAGGCTAGAAA 634  
1196 CGGGCCATTCATGGGCTCCAGGTCACTTCTACTTCAATGAAGTGGGCTGAGCCGAGG 1255  
635 TTGACCAAAACATTAATTGAAGTGGCTGATCAGATATCCATGCTTATTTGCTGTTG 694  
1256 GATATGAGAGAGAGGTCCCAACTGGTCAATGACCATATGAAATCTTTGTGACATTT 1315

QY 695 GAGCTCTCTTCTTGTGTTGAGGCAATCAATCTACTCTCTGTGTAAAG 750  
DB 1316 CCGGCTTCATGCTGGGTACTCCCAAGACCTGGTGTCTTGTGTGGAAGAAAG 1371

## RESULT 10

US-10-425-115-57116  
; Sequence 57116, Application US/10425115  
; Publication No. US20040214372A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 57116  
; LENGTH: 369  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(369)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: MRR4577\_152078C.1  
US-10-425-115-57116

Query Match 3.4%; Score 56.6; DB 21; Length 369;  
Best Local Similarity 78.2%; Pred. No. 2.4e-06;  
Matches 68; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 238 ACTAAGCTTCTTCTGCTGTCACTGCTGTTAAGTGTGATGCTGATCCTATAT 297  
DB 283 ACATTGCTTTCTCTGCACTTAATGCTGCTTAATGATGCTGCTTGAACCTATAT 342  
QY 298 CAACAGACCATCCGAGTGAAGTCA 324  
DB 343 CAGATGACCATCCGAATGCAAGTCA 369

## RESULT 11

US-10-425-115-61794  
; Sequence 61794, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 61794  
; LENGTH: 962  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRR4577\_15634C.1  
US-10-425-115-61794

Query Match 3.1%; Score 51; DB 21; Length 962;  
Best Local Similarity 80.0%; Pred. No. 0.00025;  
Matches 60; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy	1555	GTTAAGTCAATACCTTTTATGAAGACTCTGAGAGCATTTCCGGGGAAAGAAACCGCAGTC	1614
Db	32	GTGAAGTCAATACCTACTATGAGAGTCTCTAAAGCTTTCCAGATGGAACACGCACTC	91
Oy	1615	TGGGAGCTTGATATAA	1629
Db	92	TGGGAGCTCGATATAA	106

RESULT 12  
 US-10-425-115-93908  
 Sequence 93908, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 93908  
 LENGTH: 374  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_185631C.1  
 US-10-425-115-93908

Query Match	2.7%	Score	44	DB	21	Length	374
Best Local Similarity	83.3%	Pred. No.	0.021				
Matches	50	Conservative	0	Mismatches	10	Indels	0
						Gaps	0

```
Oy      321 GTACAGGCCCTGACAGTGGTCCTTTCAAGCCATCACAGAAATTGGAATCCAGSTTAACATAAACAAG 380
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       1  GTCAGAAGCCGTAGTAGTGCTCTTTCTGCAATTACAGAGCTGGACCCTGGTATATATAACAAG 60
```

```

1 RESULT 13
2 US-10-465-217-13/c
3 ; Sequence 13, Application US/10465217
4 ; Publication No. US20030204859A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Kazemi-Esfarjani, Parsa
7 ; APPLICANT: Benzner, Seymour
8 ; TITLE OF INVENTION: AN ANIMAL MODEL OF POLYGLUTAMINE
9 ; TITLE OF INVENTION: TOXICITY, METHODS OF USE, AND MODULATORS OF POLYGLUTAMINE
10 ; TITLE OF INVENTION: TOXICITY
11 ; FILE REFERENCE: 06618-686001
12 ; CURRENT APPLICATION NUMBER: US/10/465,217
13 ; CURRENT FILING DATE: 2003-06-18
14 ; PRIORITY APPLICATION NUMBER: US/09/639,207
15 ; PRIORITY FILING DATE: 2000-08-14
16 ; PRIORITY APPLICATION NUMBER: US 60/148,934
17 ; PRIORITY FILING DATE: 1999-08-12
18 ; PRIORITY APPLICATION NUMBER: US 60/148,933
19 ; PRIORITY FILING DATE: 1999-08-12
20 ; PRIORITY APPLICATION NUMBER: US 60/177,047
21 ; PRIORITY FILING DATE: 2000-01-18
22 ; PRIORITY APPLICATION NUMBER: US 60/205,720
23 ; PRIORITY FILING DATE: 2000-05-19
24 ; NUMBER OF SEQ ID NOS: 69
25 ; SOFTWARE: FastSeq for Windows Version 4.0
26 ; SEQ ID NO 13
27 ; LENGTH: 486
28 ; TYPE: DNA
29 ; ORGANISM: Artificial Sequence
30 ; FEATURE:
31 ; OTHER INFORMATION: Synthetic DNA
32 US-10-465-217-13

```

Query Match	2.4%	Score 38.8;	DB 18;	length 486;
Best Local Similarity	56.2%	Pred. No. 1.1;		
Matches	73;	Conservative	0;	Mismatches 57;
			Indels	0;
			Gaps	0

Qy	662	TTGATCAGACATTCATGCTATTTATCGCTGTGAGAGCTCTCTCTGTGCTGTAGG	721
Db	183	TTGCTGCTGCTGTGCTGCTGCTGTGCTGTGCTGCTGCTGTGCTGCTGCTG	124
Qy	722	CATCAAAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	781
Db	123	CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	64
Qy	782	GGTTGTTGTC	791
Db	63	GGGGGTTGAC	54

```

RESULT 14
US-10-719-993-6824
Sequence 6824, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CU001496
CURRENT APPLICATION NUMBER: US/10/719,993
CURRENT FILING DATE: 2003-11-24
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6824
LENGTH: 325446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)---(325446)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Table
US-10-719-993-6824

```

Query Match	2.4%	Score 38.6;	DB 21;	Length 325446;
Best Local Similarity	46.7%	Pred. No. 74;		
Matches 122;	Conservative 0;	Mismatches 139;	Indels 0;	Gaps 0

QY	520	CTCAATGACCTTTGAAACGCTTTAAACCTCGTGCGGCGCTTCGATAGATTTGGC	579
Db	142302	CTACTGTAACTTTGCACATCTTTATGATGATATGTTTTTGTCACTGTACCTCAATTTGCT	142361
QY	580	CATTCAATGGGTAAATATGCTCTTCAGATACCTTTCGAAATGCGCTAGAGGCTAGAAATTGCA	639
Db	142362	CATCTATTAATGATTGCAAAATGATGATCACTCAATCTTAATTCAACTGCTAGATCAATGTAATGATACA	142422
QY	640	CCAAAACATTATTTGAAAGTGGCTTATATGCAATATCCATGCTTAATTTCGCTGTTGGAGCT	699
Db	142422	ATGAAAAATGTTCTATTAAGGTAAAGTCTCTTTCATATGATGATCAATTTTAAAGATCC	142481
QY	700	CCCTCTCTTGGTCTCTGTTGAGGCATCAAAATCTACTCTCTGCTGTAACTTTGGCCTT	759
Db	142482	TCTCAAGGGGGGCTCTTTGTACTGAAGTGTTTAATTAACTTTAATGTAAGCATGTGCCAT	142541
QY	760	CGTGTTTCTGAGGAACTGCT	780
Db	142542	GGTGAAATCTTTGCACATTAAT	142562

RESULT IS  
US-10-424-599-56537  
; Sequence, 56537, Application US/10424599-56537  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovacic David K

APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 56537  
LENGTH: 1545  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_22062C.1  
US-10-424-599-56537

Query Match 2.3%; Score 38; DB 19; Length 1545;  
Best Local Similarity 53.3%; Pred. No. 3.9;  
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY	565	TCTATAGTATTTGCCCATTCATGGGTATATATGCTCTTCAGATACCTTCTGGAATGGCTG	624
DB	956	TCTATATTTTGGTGCAACAGTTGGTTAGTATATCTTATATACCTTGGAACAATAGCTG	1015
QY	625	AGGCTAGAAATTGCACCAAAACATTATTGAAGTGCGTTGATCAGCATATCCATGCTTAT	684
DB	1016	AATCCAGGGGTTGCTCTATATATTTTCATCTTTGGTGTCTTGTCAATCATGCTTTGGGTCT	1075
QY	685	TTGCGTGTGGAGCTCTCTTCTTGTTCT	714
DB	1076	TTGGCGTCTCATGACATCTTATATGTA	1105

Search completed: November 9, 2005, 04:41:24  
Job time : 1395 secs

**THIS PAGE BLANK (USPTO)**



Db	183	GTGGTCTGTGCTGGATGTGGCCCTTACTCGCCCTTGGACTTCAACCCCTCTGACTCCGTCTG	242
OY	228	GCTAGACACACTAAAGCTTCTTTCTGCTGTCAACTGCTGGTTTAAAGTATGTGCTTGA	287
Db	243	GCTGACACCCGCAAGCTCTTATCTGCTGTCAATTGCTGGCTCAAAATGATCTCTCGA	302
OY	288	TCCTTATATCAAAACAGACCATCCCGAGTGTAAAGTACAGGGCTGTGACAGTGTCTTACG	347
Db	303	CCCTTACAAACGAGCGGATCATCTCAGAGTGCAAAGTCAAGGCCGTGACAGTGTCTTTCG	362
OY	348	CATCAAGAAATTTGATCCAGGTTACATTAACAGGTCTCTTTTCTACTGTCTGAAAGAGTG	407
Db	363	AATTACAAAGTTAGACCTGTGTTATATACTGTGCTCTCTTCTTCAAGTGTGAAAGAAATG	422
OY	408	GCTTAAAGTGTGTGTGAGTTGTATAGAAACAAATGCAATTTGCGTGTTCATPACGA	467
Db	423	GGTCAAAATG-----	431
OY	468	TTGGAGATTTGTCAACCAACCAATTGGAAGAGCGTGACCTTATCTTCAACAAAGCTCAAGTT	527
Db	432	-----GTT	434
OY	528	GACCTTTGAAACTGCTTTAAAATCCGTTGGCGGCCCTTATAGATTTTGGCCATTCAAT	587
Db	435	AACCTTTGAAACTGATTGAAACTTCGAGAGAGGCGCTTTTATAGTTGTGCTCATTTCCAT	494
OY	588	GGGTATATATGTCTTCAGATACTTCTGGAATGTGCGGTGAGGCTAGAAATTTGACCAAAACA	647
Db	495	GGGTATATATGTGTTCGTACTTTCTGGAATGTGTTGAACATGAAATGTGCTCCCAAGA	554
OY	648	TTATTTGAAAGTGTGCTTGTATCAGCATATCATGTCTTATTTGCTGTGAGCTCTCTTCT	707
Db	555	CTACATCCGATGGCTTGAACGAACATATACATGCAATCTTGGACGTGGGCACTCTTCT	614
OY	708	TGGTCTGTGTTGAGCAATCAAAATCTACTCTCTGCTGTGAAAGTTTGGCCTTCCGTGTT	767
Db	615	TGGTTCTACTGAGACAGTTAAAGCTGCTCTTCTTGAGACACATTTGGTCTTCCAGTCAAG	674
OY	768	TGAGGGAATCTGCTGGTGTGTGTCCAATTCYTTTSCGTGTCATYTGAGCTTATGCCATT	827
Db	675	CGAGGGAAACAGCAGATGATGTGTTAAAGCATTTTGCTTATGCTTATGTGCTATCCCTT	734
OY	828	TTCAAGAATTTGCAGAGGTGATTAACAATCCGTGAGACGATTTTCTGAGGGGTGCTGCAAA	887
Db	735	CTCAGAAATTTGCCAAAGCTGATTAATATATCTGGAAGCATTTCTTTGAGGGAAGGGA--	792
OY	888	GAAAGATTAAGCGCGTATACCACTGTGATGAGAGGAAATATCAATCAAAATATTTCTGGCGTG	947
Db	793	-GGTTGCCACCAAGACAGCAATGTGATGAATGGAATATTAAGTGTGAATACTCAGGATG	851
OY	948	GCCGACAAATATTAATTAACATTGAAATTCCTTCCACTAGCGTTACAGAAACAGCTCTGT	1007
Db	852	GCCCAACAACCTTGTGCACATCTGAGAGTTCTTACAGTTGGAATGAGGAACATACCCTATC	911
OY	1008	CAACATG-----ACGACATGGAATGTGGCTTCCACACCTT	1048
Db	912	CATTATGATATCAACTGAGAACTATAACATCCAGCATGGAAGTGTGAAAGCCCACTTAT	971
OY	1047	GTCCTTTCACAGCCCGTGAATAGCAGATGGGACTCTTTTCAAAAGCAATAGAACTATGA	1106
Db	972	GTCAATTTTCTGCTAGGAGAGTTTTCAGATGGTACTCTGTTTAAAAACATTAAGGATTTGGGA	1033
OY	1107	CCGAGATGCAAGAGAGATGTTACACACAGTTAAAGAAATTTGATCATGATGACCCGTGTT	1166
Db	1032	CCCTCAGAGCATAGGGCTTATTCATCTAGCTGTGAGAGTATCAACAGGGGTGATCCAGTCT	1091
OY	1167	TAACTCTCTGACTCCTTGGAGAGAGACCACTTAATAAAATATTTTCACTATATAGTGTC	1228
Db	1092	TAACTCCCTCAACCCCTGGAGAGACCCCAATTAAGATGTGTTCTGATATATAGGCT	1155
OY	1227	TCATCTAAAGACAGAGTTGGTTATTTACTTTGCCCCAAGTGGCAAACTTATCTGTATPA	1286

Db	1152	TGATTCAAAGCTGAGGAGGCTATTATTTTGCACCGAGTGGAAAAACCATATCCAGATTA	1211
OY	1287	TTGATCATCATCGGATATCATTTACGAACCTGAAGGTTCCCTCGTGTCAAG-----	1338
Db	1312	CTGATATATATCTGATATTATTTATGAAATTTGAAGGGCTCTACTATCTCAGGCAATATA	1274
OY	1339	-----TCTGAAACTGTGTGTATGAGGAACGCTGAC	1376
Db	1272	AATAGTACACAAATTTATTTCTTCCAGATCAGGTATTTCTGTACTGSAAGGCCAACAA	1333
OY	1371	TATATCTGGGGATGAGACGGTACCCCTATCATTTCACTCTCTTGGTGCAAAATTGGCTCGG	1433
Db	1332	TTCCAGCGSAGACGGAACGGTATCTTCAATTTCTCTCATTGAGTGCAAACTGGCTTGG	1391
OY	1431	ACCTAAAGTTACATTAACAATGGCTCCCGACGCCAAGACAGATGGAAAGCGAAGTACATGT	1499
Db	1392	GCCAAAGATGAACATTAACCTAGGGCTCCGACAGCGCAAAATGATGATCTTGATTTTCAAC	1451
OY	1491	GGAACCTAAATGTTGATCATGAGCATGGGTCCAGACATCATAGCTTAACATGACAAAGCACCC	1558
Db	1452	AAAGATGATGTGCAACCAACCATGTAGGCGAGGGCATACTCCAAACATGACGAGGACCTCC	1511
OY	1551	AAGGTTAAGTACATTAACCTTTTATGAAAGCTCTGAGAGCATTTCCGGSGAAAGACACCGC	1610
Db	1512	ACATGTGAAGTACATTAACCTACTTTGAAAGATGTGAAAGCAATTCGGGATGAGAAACAGC	1574
OY	1611	AGCTGGAGGCTTATATAA	1629
Db	1572	CGTCTGGAGCTTATATAA	1590

FEATURES	LOCATION/Qualifiers	1. .596
LOCUS	AU237072	596 bp
DEFINITION	AU237072 RAF1.5 Arabidopsis thaliana clone RAF1.5-47-A08 5'	linear EST 01-APR-2002
ACCESSION	AU237072	
VERSION	AU237072	
KEYWORDS	mrna sequence.	
SOURCE	AU237072.1 GI:19876241	
ORGANISM	Arabidopsis thaliana (thale cress)	
REFERENCE	Arabidopsis thaliana Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosoids; eurosid II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 596)	
AUTHORS	Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K., Enji,A., Oono,Y., Sakurai,T., Carrinci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinozawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.	
TITLE	Large scale analysis of Arabidopsis full-length cDNA	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: maseki@rc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FL-C1 vector (Carrinci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBlueScript vector. Please visit our web site ( <a href="http://www.gsc.riken.go.jp/e/plant/index_e.html">http://www.gsc.riken.go.jp/e/plant/index_e.html</a> ) for further details.	

```
FEATURES
source
location/Qualifiers
1..596
/mol_organization="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL15-47-A08"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
```



ORIGIN /clone lib="RAF15"  
/note="Site\_1: BamHI; Site\_2: SalI"

Query Match 29.5%; Score 484; DB 1; Length 596;  
Best Local Similarity 99.4%; Pred. No. 1e-133;  
Matches 484; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAGCGCAATTCGAAATCAGTAAACGGCTTCTTCAACCGTCATCGCCGCTTTTCTTG 60  
DB 110 ATGGAGCGCAATTCGAAATCAGTAAACGGCTTCTTCAACCGTCATCGCCGCTTTTCTTG 169  
QY 61 ATTTCGCGTGGCCGAATCGCGTGAGATGAGACCGAGTTTCAACCGGCACTACTCGAAG 120  
DB 170 ATTTCGCGTGGCCGAATCGCGTGAGATGAGACCGAGTTTCAACCGGCACTACTCGAAG 229  
QY 121 CTATCGGGTATATATCATTCGCGGATTTGGCTGACGCGACCTACGACCGTGTGATCCTT 180  
DB 230 CTATCGGGTATATATCATTCGCGGATTTGGCTGACGCGACCTACGACCGTGTGATCCTT 289  
QY 181 GACTGCTCATACACTCCGTTGAGACTTGCATCGCTGACCTGATGCTAGACCACT 240  
DB 230 GACTGCTCATACACTCCGTTGAGACTTGCATCGCTGACCTGATGCTAGACCACT 349  
QY 241 AAGCTTCTTCTGCTGCTCAACTGCTGTTAAGTATGATGCTAGATCCTATATATCA 300  
DB 350 AAGCTTCTTCTGCTGCTCAACTGCTGTTAAGTATGATGCTAGATCCTATATATCA 409  
QY 301 ACAGACCATCCCGAGTGAAGTACAGCGCCGTCAGAGTGTCTTTCAGCCATCACAGATTG 360  
DB 410 ACAGACCATCCCGAGTGAAGTACAGCGCCGTCAGAGTGTCTTTCAGCCATCACAGATTG 469  
QY 361 GATCCAGGTTATATACAGAGCTCTTCTTCTACTGCTGGAAGAGTGGCTTAAGTGTCT 420  
DB 470 GATCCAGGTTATATACAGAGCTCTTCTTCTACTGCTGGAAGAGTGGCTTAAGTGTCT 529  
QY 421 GTTGAGTTGATATGAGCAATGCAATTTGCGCTGTCATAGATGAGATGATGCA 480  
DB 530 GTTGAGTTGATATGAGCAATGCAATTTGCGCTGTCATAGATGAGATGATGCA 589  
QY 481 CCAACCA 487  
DB 590 CCAACCA 596

RESULT 3  
LOCUS CK259394 946 bp mRNA linear EST 03-AUG-2004  
DEFINITION EST05472 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB029 5' end, mRNA sequence.  
ACCESSION CK259394  
VERSION CK259394.1 GI:39816372  
KEYWORDS Solanum tuberosum (potato)  
SOURCE Solanum tuberosum  
ORGANISM  
EST.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 946)  
Buell, C.R., Hart, A., Ziemann, V., Karameycheva, S.A. and Baker, B.  
Generation of ESTs from abiotic stressed potato tissue  
Unpublished (2003)  
Other ESTs: EST705470 EST705471 EST705473  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: CAG GAA ACA GCT ATG ACC.  
Location/Qualifiers  
1..946  
/organism="Solanum tuberosum"

/mol type="mRNA"  
/culti var="kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB029"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab host="DH10B-TorA"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

## ORIGIN

Query Match 28.3%; Score 464.2; DB 7; Length 946;  
Best Local Similarity 70.7%; Pred. No. 1.1e-127;

Matches 671; Conservative 0; Mismatches 253; Indels 25; Gaps 3;

QY 678 TCCCTATTTGCTGAGAGCTCCCTCTTGTGTTGAGGCAATCAATCTACTCT 737  
DB 1 TCCCTATTTGCTGAGAGCTCCCTCTTGTGTTGAGGCAATCAATCTACTCT 60  
QY 738 CTCTGCTGTAAGTGTGCTTCTCTGTTTCTGAGGAACTGCTGCTGTTGTCATTC 797  
DB 61 ATCGGGGCTCATATTTGCTTCTCTGTTTCAAGAGGAGCAGCTCCCTGATGTTCAATTC 120  
QY 798 TTTTGGCTGCTGATGCTGCTTATGCAATTTTCAAGAGGCTGATGATGCAATTC 857  
DB 121 ATTTGATCTTCTATATGATGTTTCAATTTTCAAGAGGCTGATGATGCAATTC 180  
QY 858 CTGAGCGATTTTCTGAGGAGTGTGCAAGAAAGATAGCGCTATACCACTGTGATGA 917  
DB 181 TCGAGACATTTCTCT---GGTGAACAGAAATATCCATCATGCTTACCTTGGATGA 237  
QY 918 AGAGGATATCATCAATATATTTCTGCTGCGCCCAATATATTTAACTTGAATTTCC 977  
DB 238 GCATGATCTTATATGATGATCTGCTGCGCCCAATATATTTGAATTTCC 297  
QY 978 TTCCACTAG-----CGTTACAGAAACGCTCTTATGTAACATGAC 1016  
DB 238 TTCCACTAG-----CGTTACAGAAACGCTCTTATGTAACATGAC 357  
QY 1017 CAGCATGAGATGAGCTTCCACCTTTTGTCTTTCACAGCCGCTGATGAGATGAG 1076  
DB 358 TCGAATGAGATGAGATTTCCGACCAATTTGCTCTTCTGCTGTAAGTCTTGACGG 417  
QY 1077 GACTCTTTTCAAGCAATGAGAGACTATGACCCAGATGACAGAGATGTTACACAGTT 1136  
DB 418 AACCTTTTCAAGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATG 477  
QY 1137 AAAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1196  
DB 478 AAAAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537  
QY 1197 TATATAAATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1256

Db 538 TCTCAAGACATATCTGCATTATAGCCATGAGTGGTGGTACTATT 597  
1257 TCCCCAAGTGGCAACCTTATCTGATATGATCATCAGCATATCATTAAGAAC 1316  
Db 598 TGCACCAAGTGGGAACCTTATCTGATATGATCATCAGCATATCATTAAGAAC 657  
Qy 1317 TGAAGTTCCTCGTGTCAAGGTCTGGAACCTGTGTGATGGGAACCTGACCTATAAC 1376  
Db 658 CCAAGGATCTGTATATCCCGGTCAAGGAATCTGGTTGAAGGAACCTGTGTCAAG 717  
Qy 1377 TGGGATGAGACGATACCTATCATCTCTTGTGTGCAAGATTTGGCTGGAACCTTA 1436  
Db 718 TGGGATGAGACGATACCTATCATCTCTTGTGTGCAAGATTTGGCTGGAACCTTA 777  
Qy 1437 AGTTAACAATCAATGAGTCTCCCAAGCAGACACGATGAGAGAGTATGAGAACT 1496  
Db 778 AGTGAACAATCAAGACACACAGTCAAGATGATGATGATGATGATGATGATGAT 837  
Qy 1497 AATGTTGATCAT-GAGCATGGGTGACAGATCATAGCTAATGATGATGATGATGAT 1555  
Db 838 AATATAGACATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 897  
Qy 1556 TTAATGATCAATCTTTTATGAGACTGTGAGAGATTTCCGGGAAGAG 1604  
Db 898 TGAAGTACATCACTATTATGAGATTTGGAAGTTTCCAGGAACGAG 946

RESULT 4  
LOCUS CF844688 752 bp mRNA linear EST 30-OCT-2003  
DEFINITION psbH029xf13f USDA-IPAFS: Expression of Phytophthora sojae genes during infection and propagation\_sbh Phytophthora sojae cDNA clone sbH029f13 5, mRNA sequence.  
ACCESSION CF844688  
VERSION CF844688.1 GI:38060342  
KEYWORDS EST.  
SOURCE Phytophthora sojae  
ORGANISM Phytophthora sojae  
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.  
REFERENCE 1 (bases 1 to 752)  
AUTHORS Tyler, B. Not Published  
TITLE Unpublished (2003)  
JOURNAL Contact: Tyler B  
COMMENT Tyler lab  
VBI  
1880 Prate Dr., Blacksburg, VA 24061, USA  
Tel: 540-231-7318  
Email: bmtyley@vt.edu  
PCR Primers  
FORWARD: BK reverse primer  
BACKWARD: BK reverse primer  
Plate: 029 row: F column: 13  
Seq primer: BK reverse primer  
High quality sequence stop: 752.  
Location/Qualifiers  
1..752  
/organism="Phytophthora sojae"  
/mol\_type="mRNA"  
/db\_xref="taxon:67593"  
/clone="sbH029f13"  
/cissue\_type="mycelium"  
/cell\_line="P6497"  
/dev\_stage="48 hr. post infection stage"  
/lab\_host="Soybean plant"  
/clone\_lib="USDA-IPAFS: Expression of Phytophthora sojae genes during infection and propagation\_sbh"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 26.7%; Score 438.8; DB 7; Length 752;

Best Local Similarity 79.9%; Pred. No. 4,5e-120;  
Matches 526; Conservative 0; Mismatches 131; Indels 1; Gaps 1;  
Qy 108 CGACATACGAGAGCATGCGGATATATCATTCGCGGATTTGCGTCAAGCAGCATACGAGC 167  
Db 94 GCACTACTGAGAGCTCTCCGATATATATCCCTGCGTTCGATCATTCAAGCTCCAGC 153  
Qy 168 GTGTGATCTCTTGAACCTGATACACTCCGTTGAGCTTCAATCCGCTGACCTGTATG 227  
Db 154 ATGGTCATATTCGACCTGTCTTACTCTCCGCTGATTTCAACCTCTCGATTTGGCTG 213  
Qy 228 GCTAAGACCACTAAGCTTTCTTCTGCTGTCAACTGCTGTGTTAAGTATAGGCTTGA 287  
Db 214 GCTCAGCAACAACCAACTCTTCTGCTGTCAATGCTGCGCTTAAGTATAGGCTTGA 273  
Qy 288 TCCATATATCAACAGACATCCCGAGTATGATGACAGGCTGACAGTGGCTTTGACG 347  
Db 274 TCCATATCAATCAAGACATCAATCCGATTTGATGAGATTCCTGATAGGCTTTCTG 333  
Qy 348 CATCAACAATGATTCAGGATTAATCAAGATCTCTTCTACTGTCTGGAAGAGTG 407  
Db 334 TATTACAGAACTGATCCGGTTATATTAAGAGACCTCTTCATCGGTTGGAAGATG 393  
Qy 408 GCTTAAGTGTGTGTGAGTTTGTGATTAAGCAATGCTATGCTGCTTTCCATAGA 467  
Db 394 GATTAAAGTGTGTATTAATTTGATTAAGCAATGCAATTAATGCTGTTCATATGA 453  
Qy 468 TTGAGATTTGTCAACCAACCAATGGAAGAGCGTACCTTACTTTCAAGCTCAAGT 527  
Db 454 TTGAGATTTGTCCCTTCAAGCTTGAAGAGCGGACCTTACTTTCAATGATTAAT 513  
Qy 528 GACCTTGAACCTGTTAAACTCCGCGCCCTTCTATAGATTTGCCCATTCAT 587  
Db 514 AACATTTGAACCTGTTCAAACTTCGCGGCGCCCTTCTATAGTTTGGCCATTCAT 573  
Qy 588 GGGTAAATATCTTCTGATCTTTCGATATGCTGAGCTGAGGCTTAAGAAATGGAACCA 647  
Db 574 GGGTAAATATCTTCTGATCTTTCGATATGCTGAGCTGAGGCTTAAGAAATGGAACCA 633  
Qy 648 TTATTTGAAGAGGCTT-GATCAGATATCAATGCTTATTTGCTGTGAGACCTCTTTC 706  
Db 634 TTATTCAGAGGCTGTGATCAATATTCGCTCTATNTTGTCTGTGAGACCTCTTTC 693  
Qy 707 TTGCTTCTGTGAGCAATCAATCTCTCTGTGTGATGAGCTTGGCTTCTCTGT 764  
Db 694 TTGCTTCTGTGAGCAATCAATCTCTCTGTGTGATGAGCTTGGCTTCTCTGT 751

RESULT 5  
LOCUS CK939714 713 bp mRNA linear EST 19-MAR-2004  
DEFINITION CGF1004746 F04 Developing fruit flavedo at 165 DAFB Citrus sinensis cDNA clone\_F1650002\_1IF\_F04 5', mRNA sequence.  
ACCESSION CK939714  
VERSION CK939714.1 GI:45452970  
KEYWORDS EST.  
SOURCE Citrus sinensis  
ORGANISM Citrus sinensis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Sapindales; Rutaceae; Citrus.  
REFERENCE 1 (bases 1 to 713)  
AUTHORS Urasu, S., Baek, J., Leslie, A., Xu, J., Cook, D. and Dandekar, A.  
TITLE Analysis of peel specific genes in Citrus (2004)  
JOURNAL Unpublished (2004)  
COMMENT Contact: Abhaya Dandekar, PhD  
CAES Genome Facility  
UC Davis, Department of Pomology  
One Shields Ave, Davis, CA 95616, USA  
Tel: 530 752 7784  
Fax: 530 752 8502  
Email: amandekar@ucdavis.edu  
Seq primer: WSC-F-TCGAGATCTGAGCAGC.

FEATURES  
source

Location/Qualifiers  
1.713  
/organism="Citrus sinensis"  
/mol\_type="mRNA"  
/cultivar="Washington navel orange"  
/db\_xref="taxon:2711"  
/clone="F1650002 JIF\_F04"  
/sex="Hermaphrodite"  
/cell\_type="Palvado"  
/dev\_stage="Developing fruit sample-collected October 22, 2003"  
/lab\_host="X110-Gold"  
/clone\_lib="Developing fruit flavedo at 165 DAFB"  
/note="Organ: Fruit; Vector: pT1p1ex2; Site: 1: SflrA; Site 2: SflrB; Developing citrus fruits were harvested from trees growing in the 'Citrus variety collection' in the Wolfkill experimental orchard located in Winters, California (USA). Fruit was collected on October 22, 2003, between 8 to 9 am and stored at 4C. The flavedo tissue was dissected out of developing fruit (165 DAFB) and used to isolate RNA using Trizol reagent from Invitrogen. The cDNA library was constructed using the SMART cDNA library kit (Clontech). The primary library was an masee evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (XL10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/tray), picked using a Qbot and archived in 384 well dishes."

## ORIGIN

Query Match 24.2%; Score 397; DB 7; Length 713;

Best Local Similarity 76.9%; Pred. No. 1.7e-107; Matches 484; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

28 GCTTCCTTACCGTCATCCCGCTTTTCTTGATTTGGCGTGGCCGAAGCGGTGAG 87  
84 GCTGTGCT 143  
88 GATGAGACCGAGTTTTCACGCGCACTACTGGAAGTATCGGGTATTAATCTCCGGATT 147  
144 AAGGCGCGGAGTTTCAACCGGCGGCTACCCCAAGTTGTCGGATATTAATACAGGCTTT 203  
148 GGGTGACGACCTACGACGCGGTGATCTTGAATGTCATACACTCCGTGACCTTC 207  
204 GGGTGACGACCTGAGACGATGATTAATCTGACGTGTCGACTCTCTCTCGATTC 263  
208 AATCGGCTGACCTCGTATGCTAGACACCACTAAGCTTTCTTCTGCTCAACTGCTG 267  
264 AACCTCTGACCTCGTCTGCTAGACACTACCAACTCTTCTCTCTGAACTGTGG 323  
268 TTTAAGTATGCTGCTAGATCTTATTAATCAACAGACCATCCGAGTGTAAATGACG 327  
324 CTTAAGTATGATACACTAGATCTTATCAATCAACAGATTAATCCCAATGCAAGTCTCA 383  
328 CCGTACAGTGTCTTTCAGCATACAGAAATGATCCAGTTACATAAGTCTCTT 387  
384 CCGTATAGTGTCTTTCGCAATTACAGAACTTGAATGATTAATTAAGTCTCTG 443  
388 TCTACTGTGGAAGAAGTGGCTTAAAGTGTGTTGAGTTTGTATAGAACCAATGCA 447  
444 TCTCTAGTGTGGAAGAAGTGGTAAAGTGTGTTGAGTTTGTATAGAGGTATATCA 503  
448 ATTGTGCTGTTTCCATACGATTGGAGATTGTCAACCAATTTGGAAGAGCTGACTT 507  
504 ATCATTTGCCGCTCCATATGATGAGATTGTCAACATCAAGCTCCAGAGAGAGACTT 563  
508 TACTTTCACAGCTCAAGTTGACTTGAACGCTTTAAATCCGCTGGCGGCGCTTCT 567  
564 TATTTTCAAACTAAAGTTGATGAACTGCCCTTAATCTGTGTGAGCCCTTCA 623  
568 ATAGTATTTGCCCATTCATGAGTAAATGCTTTCAGATCTTTTGAATAGGCTGAG 627

Db 624 TTAGTTCCTGCTCATTCATGAGTAAATATGTTTCTGACTTCTGAGTGGCTAAG 683  
Qy 628 CTAGAAATTGCACCAAAACATTATTGAA 656  
Db 684 TTAGAAATTCCCAAAACATATATCA 712

RESULT 6  
CK259393  
LOCUS  
DEFINITION  
EST705471 potato abiotic stress cDNA library Solanum tuberosum cDNA clone FOAB029 5' end, mRNA sequence.  
ACCESSION  
CK259393  
VERSION  
CK259393.1 GI:39816371  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Solanum tuberosum (potato)  
Solanum tuberosum  
Buckeye; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE  
1 (bases 1 to 788)  
Buell,C.R., Hart,A., Ziemann,V., Karamycheva,S.A. and Baker,B.  
AUTHORS  
Generation of ESTs from abiotic stressed potato tissue  
JOURNAL  
Unpublished (2003)  
COMMENT  
Other ESTs: EST705470 EST705472 EST705473  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/

Seq primer: ART TAG CTG ACA CTA TAG.  
Location/Qualifiers

FEATURES  
source

1.788  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="FOAB029"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

## ORIGIN

Query Match 23.5%; Score 386.2; DB 7; Length 788;

Best Local Similarity 71.2%; Pred. No. 3.1e-104; Matches 563; Conservative 0; Mismatches 203; Indels 25; Gaps 3;

Qy 678 TCGTATTTGCTGTGAGCTCTCTTGTGTTGAGGCAATCAATCTACTCT 737  
|||||

Db 1 TGCCTATTTTGTCTGTGAGCTCTCTCTGTGCTATGAGACATCAACACATTT 60  
Qy 728 CTCTGGTGT-AAAGTTTGGCCTTCCCTGTTCTGAGGGAACCTGCTGGTGTGTTCAATT 796  
Db 61 ATCCGGGCTCTACATTTTGTCTCTCTGTTTCAGAGGGACAGCTCCCTGATGTTCAATA 120  
Qy 797 CTTTTCGGTCTGATTTGGCTTATGCAATTTTCAAGATTGCAAGGGTGAATACACAT 856  
Db 121 CATTTGATCTCTATATGATGATTACATTTTCAAAATATTTGTAACAATGATTAATGAT 180  
Qy 857 CCTGACGATTTTCTGGGGGTCTGCAAGAAAGATTAAGCCGATATCCACTGTGATG 916  
Db 181 ATCGAGAGATTTCTCT---GGTGAAGACAGAAATACATATGCTTACATTTGAGATG 237  
Qy 917 AAGAGAAATATCAATCAAAATATTTCTGCTGGCCGCAATATATTATTAATTGAAATTC 976  
Db 238 AGCATGAACTTAAATCAAGTACTCTGCTGGCCAAACATATCAATATTGAAGTTTC 297  
Qy 977 CTTTCACTAG-----CGTTACAGAAACAGCTCTAGTCAACATGA 1015  
Db 298 CTTTCATTCGAGGGAATGAGTTTATCCCTCAGATTGTAGAAACCTGCCAAACCACTTGT 357  
Qy 1016 CCAGATGGAATGTGGCCTTCCACCCCTTTGTCTTTCACAGCCGCTGAACATGACATG 1075  
Db 358 CTGGAATGGAATGTGGATTCGACACAAATGTCTTTCTGCTGTAAGTCTCTGACG 417  
Qy 1076 GGAATCTTTTCAAGCAATAGAAAGTATGACCCAGATAGCAAGATGTTACACAGT 1135  
Db 418 GAACCTTTTTCAGAGCAATTAAGAAATTAATGATCTGACAGTGAAAGACTTTTCACTTT 477  
Qy 1136 TAAAGAGTTGATCATGATGACCTGTTTAAATCTCTGATCTGATCTGTTGGAGAGACAC 1195  
Db 478 TAAAAAGCTTACATGATGATCGATTTTAAATCTCTAACACCTTGGAGAAAGCACAC 537  
Qy 1196 CTATATAAAATGATTTTGCATATATGTCATCTATTAAGACAGAGTTGGTATTACT 1255  
Db 538 CTCTCAAGAAATATTTCTGATTTATGCGTGAATTCGAAAGCTGAGTGTGTTACTATT 597  
Qy 1256 TTGCCCCAAGTGGCAAACTTATCTGATTAATTTGATCATCAAGATCATTTACGAA 1315  
Db 598 TTGCACCAAGTGGGAAGCTTATCTGATTAATTTGATTAATTAACAGATGATATATGAA 657  
Qy 1316 CTGAAGTTCCCTCGTCAAGGTCTGAACTGTGGTTGATGAGAAAGCTGTGACCTATTA 1375  
Db 658 TCGAAGGATCTCTGATTTCCCGGTCAAGAAATCTGGTTGAAGAAACCTGTGTGCAACA 717  
Qy 1376 CTGGGAGTGAAGCGGATACCTTATCATCTCTCTTGGTGAAGAAATTTGGCTCGACCTA 1435  
Db 718 GTGGGAGTGAAGCGGATACCTTATCCCTCTCTGTTGCAAGAAATTTGGCTGGGAGCA 777  
Qy 1436 AAGTTAACATA 1446  
Db 778 AAGTGAACATA 788

RESULT 7  
CK260764 968 bp mRNA linear EST 03-AUG-2004  
LOCUS CK260764  
DEFINITION EST006442 potato abiotic stress cDNA library Solanum tuberosum cDNA  
clone POAB552 5' end, mRNA sequence.  
ACCESSION CK260764  
VERSION CK260764.1 GI:39817742  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
REFERENCE  
AUTHORS Bull, C.R., Hart, A., Ziemann, V., Karaymicheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST706843

Contact: Robin Buel  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@igf.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ANT TAG GTG ACA CTA TAG.  
Location/Qualifiers

FEATURES  
source  
1..968  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cuiovar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POAB552"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="RH10B-Tona"  
/clone\_lib="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: Solanum tuberosum var. Kennebec plants were  
grown from cuttings on a 16hr light/8 hr dark cycle at 25  
C for 3-4 weeks. Abiotic stress conditions were applied to  
four separate sets of plants. Set 1 involved saturation of  
the soil with 150 mM NaCl and tissues were harvested at  
following application of the salt stress (leaves: 2hr,  
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).  
Set 2 were grown under the standard conditions and then  
were water stressed by withdrawal of further watering  
applications. Drought stressed plants were harvested after  
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d  
and 5d). Set 3 were grown under the standard conditions  
and then were cold stressed by placement at 4 C. Cold  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d. Set 4 were grown under the standard conditions and  
then were heat stressed by placement at 35 C. Heat  
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,  
2d and 4d and heat-stressed roots were harvested at 6 hr,  
12 hr, 1 d, and 4d. RNA was isolated from all tissues and  
equal RNA from each tissue and stress was pooled to  
construct the cDNA library. RNA sample."

ORIGIN  
Query Match 23.0%; Score 376.8; DB 7; Length 968;  
Best Local Similarity 74.1%; Pred. No. 2.3e-101;  
Matches 477; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 988 GTTACAGAAACGCTCTAGTCAATGACAGCATGAAATGGCTTCCACCTTTTG 1047  
Db 22 GTTGAAGAACTGCCCAACCACTTGTCTGAAATGATGTGATTCCTCGACAAATTG 81  
Qy 1048 TCTTTCAGAGCCCGTGAAGTGAAGTGTATCATGATGACCTGTGTTT 1167  
Db 82 TCTTTTCTGCTCGGAAGTCTCTGATGAAACCTTTTCAAGCATTAAGATTAATGAT 141  
Qy 1108 CCAGATGCAAGAGATGTTACACCAAGTTAAAGATGTATCATGATGACCTGTGTTT 1167  
Db 142 CCGACAGTGAAGACTCTTTCACCTTTTAAAAAGTCTTACATGATGATCCGATTTTA 201  
Qy 1168 AATCTCTGACTCTTGGGAGAGACCACTTAAATAATGATTTTGCATATATGTGCT 1227  
Db 202 AATCTCTTACACCTTGGGAAAGACCACTTCAAGAACATATTTCTGATTTAAGGCTTA 261  
Qy 1228 CATCTAAGACAGAGTGTATTACTTTCCTCCCAATGGCAACCTTATCTGTATAT 1287  
Db 262 GATTGGAAGACTGAGTGTGTTACTATTTTGCACCAAGTGGGAAGCTTATCTGTATAT 321  
Qy 1288 TGGATCATCAGGATATCATTTACGAACGTAAAGTTTCTCTGTCAAGGTCTGGAAT 1347  
Db 322 TGGATTATTAACAGATGATATGAATTCGAAGGATCTGTATTTCCGGGTGAGGAAT 381  
Qy 1348 GTGCTGATGGGAAGCGCTGACCTATTAATCTGGGATGAGAGGTACCTATCATTCATC 1407  
Db 382 CTGTTGAAGAAACCTGTGTGCAACAGTGGGATGAGAGCGTGCATATCCATTCCTTC 441

QY	1408	TCTTGGAGCAAGATTTGGCTCGAAGCTTAAAGTAACTAATGAGTGGCTCCGACCGAGA	1467
Db	442	TCTCGTGCAGAAATTTGGCTCGGCGCCAAAGTGAACATACAGACACCACTGACAG	501
QY	1468	CACGATGGAACGACGATACATGTGGAACCTAAATGTGATCATGACATGGGTGCACATC	1527
Db	502	CATATATGGCTCAGCTGTACAGATGCATCACTCAAAATTTAGAGCATCAACATGTGAAGATATC	561
QY	1528	ATTACTTAACATGACAAAGACCAAGGTTTAAGTACATACCTTTATAGAACTCTGAG	1587
Db	562	ATTCCTCAATATGACAAAGTTTGTCTGACATGAAGTACATTAACCTTATTAAGAAATTTGAA	621
QY	1588	AGCATTTCCGGGGAAGAGAACCGCAGCTCTGGAGCTTGATAAAG	1631
Db	622	AGTTTCCAGAACGAGAACGACGACTTTGGAGCTTGATTAAGG	665
RESULT 8			
LOCUS	B0508623		
DEFINITION	B0508623	686 bp	mRNA linear EST 07-MAR-2003
ACCESSION	EST616038	Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMGX37	
VERSION	B0508623	5' end, mRNA sequence.	
KEYWORDS	B0508623	.2 GI:21924395	
SOURCE	EST.		
ORGANISM	Solanum tuberosum (potato)		
REFERENCE	Solanum tuberosum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.		
TITLE	Buell, C.R., Hart, A., Baker, B., Tinkaley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J., and Karamycheva, S.A.		
JOURNAL	Generation of a set of potato cDNA clones for microarray analyses Unpublished (2002)		
COMMENT	On Jun 10, 2002 this sequence version replaced gi:21367492. Other ESTs: EST616039		
	Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seq primer: T3.		
FEATURES	Location/Qualifiers		
source	1..686		
	/organism="Solanum tuberosum"		
	/mol_type="mRNA"		
	/cultivar="Kennebec or Binje"		
	/db_xref="taxon:4113"		
	/clone="STMGX37"		
	/tissue_type="mixed tissues"		
	/lab_host="SOLR"		
	/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"		
	/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Combination of untreated and Phytoptophora infestans-treated explants of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."		
ORIGIN			
Query Match	22.8%;	Score 374.8;	DB 5; Length 686;
Best Local Similarity	74.0%;	Pred. No. 8e-101;	
Matches	475;	Conservative 0;	Mismatches 167; Indels 0; Gaps 0;
QY	988	GTTACAGAAACGCTTATGATCAACATGACCGACGATGAAATGTGGCTTCCACCTTTTG	1047
Db	4	GTTGTAGAAATGCGCCAAACCAACTGTCTGTGAATGTGATTCCTGACACCAATTTG	63

Qy	1048	TCCTTTCACAGCCCGGTGAACATAGCAGATGGGACCTCTTTCAAGAGGAATAGAGACTATGAC	1107
Db	64	TCCTTTTCTGCTGCTGTAAGTCTCTGACGGAACTTTTTCAAAGCCATMAAGAAATTATGAT	123
Qy	1108	CCAGATGACGAAGAGATGTTACACCACTGTAAGAAGTTGATCATGATGACCCCTGTTTTT	1167
Db	124	CCTGACAGTGAAGAAGACTCTTTTCACTTTTAAAAAAGCTTACCATGATGATCCGATTTTT	183
Qy	1168	AATCCTCTGACTCTCTGGGAGAGACCACTATAAAAATATATTTTSCATATATNGTGCT	1227
Db	184	AATCCTCTTAACACCTTGGGAAAGCACTCTCAAGAAACATATTCGACTATTATGCGGTA	243
Qy	1228	CATCTAAAGAACAAGTGTGGTATTACTTTTGCCCAAGTGGCAAACTTATCCTGATAAT	1287
Db	244	GATTCGAAGACTGAGGTGTGGTACTATTTTGGACCAAGTGGGAAGCTTATCTGATAT	303
Qy	1288	TGATCATCAACGATATCATTTACGAACCTGAAGGTTCCCTCGTCAAGGTCTGAACT	1347
Db	304	TGATATTAAACAGATGTGATATATGAATCGAAGGATCTCTGATTTCCGGTCAAGAAAT	363
Qy	1348	GTGTTATATGGGAACGCTGGAACCTATATACCTGGGATGAGACCGTATACCTTATCACTC	1407
Db	364	CTGGTTAAAGAAACCCCTGGTCAACAAGTGGGATGACGGGTCCATACATTTCCCTC	423
Qy	1408	TCTTGTGCAAGAAATTGGCTCGACCTTAAAGTTAACATTAACAATGGCTCCCAAGCAGAA	1467
Db	424	TCTGTGTGCAAGAAATTGGCTGGGCCCAAAAGTGAACATTACAGAGAACACACAGTCAAG	483
Qy	1468	CACAGTGAAGCGACGTACATGTGGAACTAAATGTTGATCATGACATGGGTCAACATC	1527
Db	484	CATGATGGCTCAGCTGTACAAAGTCACTCAAAATATAGACATCAACATGGTGGAGATATC	543
Qy	1528	ATTACTTAACATGACAAAGAACAAGGTTAATGATACATACTTTTATGAAGACTCGAG	1587
Db	544	ATTCCCAATATGACAAAGTTAGCTACATGATGAAGATCAATTAATATGAGGATTCGAA	603
Qy	1588	AGCAATTCGGGGAGAGAACCGCAGCTGTGGAGCTTGATATAA	1629
Db	604	AGTTTTCAGAGAACGAGAACGAGCTTTGGGAGCTTATATAA	645
RESULT 9			
LOCUS	BG600021	792 bp	mRNA linear EST 07-MAR-2003
DEFINITION	EST504916 csts Solanum tuberosum cDNA clone csts27C13 5' sequence,		
ACCESSION	BG600021		
VERSION	BG600021.1	GI:13617157	
KEYWORDS	EST.		
SOURCE	Solanum tuberosum (potato)		
ORGANISM	Solanum tuberosum		
REFERENCE	Eukariota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.		
AUTHORS	1 (bases 1 to 792) van der Hoeven R., Bezzerides, J., Sun, H., Cho, J., Chiemiango, A., Bougrier, O., Buell, C. R., Ronning, C., Tanksley, S. and Baker, B.		
TITLE	Generations of ESTs from sprouting potato eyes		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seq primer: M13f-R.		
FEATURES	location/Qualifiers		
SOURCE	1..792		
	/organism="Solanum tuberosum"		
	/mol_type="mRNA"		
	/cultivar="Kennedec"		

/db\_xref="taxon:4113"  
 /clone="CSTS27C13"  
 /issue\_type="sprouting eyes from tubers"  
 /dev\_stage="12-14 weeks post harvest"  
 /lab\_host="SOLR"  
 /clone\_1ib="CSTS"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

## ORIGIN

Query Match 22.4%; Score 367; DB 4; Length 792;  
 Best Local Similarity 74.2%; Pred. No. 1.9e-98;  
 Matches 477; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

QY 988 GTTACAGAAACAGCTCTACTCAATGACACAGCATGGAATGCGCTTCCACCTTTTG 1047  
 Db 4 GTTACAGAAACAGCTCTACTCAATGACACAGCATGGAATGCGCTTCCACCTTTTG 63  
 QY 1048 TCTTTCAGACCCGCGTAATAGAGATGGGACTCTTTTCAAGCAATAGAGACTATGAC 1107  
 Db 64 TCTTTCAGACCCGCGTAATAGAGATGGGACTCTTTTCAAGCAATAGAGACTATGAC 123  
 QY 1108 CCAGATAGCAGAGAGATGTTACACCACTTAAAGATTGTATCATGATGACCTGTTTTT 1167  
 Db 124 CCTGACAGTGAAGAGACTCTTTCACCTTTTAAAGATTGTATCATGATGACCTGTTTTT 183  
 QY 1168 AATCTCTGACTCTCTTGGAGAGACCACTTATTAAGATTGTATCATGATGACCTGTTTTT 1227  
 Db 184 AATCTCTGACTCTCTTGGAGAGACCACTTATTAAGATTGTATCATGATGACCTGTTTTT 243  
 QY 1228 CATCTAAGACAGAGGTTGTTATTTACTTTGCCCCAAGTGGCAACCTATCCTGATAT 1287  
 Db 244 GATTGGAAGACTGAGGTTGTTATTTACTTTGCCCCAAGTGGCAACCTATCCTGATAT 303  
 QY 1288 TGGATCATCAGGATATCATTTACGAACTGAAGTTCCCTGATGATGATGATGATGAT 1347  
 Db 304 TGGATCATCAGGATATCATTTACGAACTGAAGTTCCCTGATGATGATGATGATGATGAT 363  
 QY 1348 GTGGTTGATGGAGAGCGCTGGACCTATTACTGGGATGAGACGATACCTATCATTCATC 1407  
 Db 364 GTGGTTGATGGAGAGCGCTGGACCTATTACTGGGATGAGACGATACCTATCATTCATC 423  
 QY 1408 TCTTGTGCAAGATTTGGCTGGACCTAAGTTAATCAATGATGATGATGATGATGATGAT 1467  
 Db 424 TCTTGTGCAAGATTTGGCTGGACCTAAGTTAATCAATGATGATGATGATGATGATGATGAT 483  
 QY 1468 CACGATGGAAGCAGATCATGATGGAATGGAATGTTGATCATGATGATGATGATGATGAT 1526  
 Db 484 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543  
 QY 1527 CATAGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586  
 Db 544 CATAGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603  
 QY 1587 GAGCAATCCGGGAGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629  
 Db 604 AAGTTTCCAGAGACGAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 646

RESULT 10  
 CVA78732  
 LOCUS CVA78732 694 bp mRNA linear EST 01-OCT-2004  
 DEFINITION Lambda Plax fiber enriched Lambda ZapII Library Linnm us1cat18s1num  
 CDNA 3', mRNA sequence.  
 ACCESSION CVA78732  
 VERSION CVA78732.1 GI:53703508  
 KEYWORDS EST.  
 SOURCE Linnm us1cat18s1num (Plax)  
 ORGANISM Linnm us1cat18s1num

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Contact: Neutelings G  
 Laboratoire de Physiologie des Parois Vegetales  
 Universite des Sciences et Technologies de Lille  
 LPPV, SN2 2e etage, USTL, Villeneuve d'Ascq F-59655, France  
 Tel: (33) (0) 3 20 43 40 29  
 Fax: (33) (0) 3 20 33 63 02  
 Email: godfrey.neutelings@univ-lille1.fr  
 Seq primer: 73.  
 Location/Qualifiers  
 1..694  
 /organism="Linnm us1cat18s1num"  
 /mol\_type="mRNA"  
 /cultiVar="Hermes"  
 /db\_xref="taxon:4006"  
 /issue\_type="Outer fiber-bearing tissues"  
 /dev\_stage="Mid-flowering stage"  
 /clone\_1ib="Plax fiber enriched Lambda ZapII Library"  
 /note="Site\_1: XhoI; Site\_2: EcoRI. mRNA were extracted from field grown plants (Dunkerque, France). Following harvest, roots, leaves and flowers were cut-off and discarded, as were the bottom 10 cm of the stem. The following bottom 15-cm section of the stem was then isolated and the outer fiber-bearing tissues peeled off and immediately frozen in liquid nitrogen. Library was constructed using ZAP-cDNA synthesis kit and ZAP-cDNA Gigapack III cloning kit (Stratagene)."

## ORIGIN

Query Match 21.7%; Score 355.4; DB 7; Length 694;  
 Best Local Similarity 75.1%; Pred. No. 5.5e-95;  
 Matches 443; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 1040 CCTTTTGTCTTTCACAGCCCGTGAATGACAGATGGAAGCTTTTCAAGCAATAGAG 1099  
 Db 1 CTCAGTATCTTTTCACAGCTCGTGAAGTTGATGAGAACGCTTTTCAGAGCATAGAG 60  
 QY 1100 ACTATGACCCAGATGACAGAGAGATTTACACCACTTAAAGATTGATCATGATGAC 1159  
 Db 61 ATTATGATTCGACAGCAAGAGGCTTTTATTCATTAAGAGGATACCATGATGATGATGAT 120  
 QY 1160 CTGTTTAAATCTCTGACTCCTTGGAGAGACCACTATTAAGATTGATTTGATAT 1219  
 Db 121 CTGTTTAAATCTCTGACTCCTTGGAGAGACCACTATTAAGATTGATTTGATAT 180  
 QY 1220 ATGTGCTATCTTAAGAAGAGGTTGTTATTTATCTTTGCCCAAGTGGCAACCTATC 1279  
 Db 181 ATGTGCTATCTTAAGAAGAGGTTGTTATTTATCTTTGCCCAAGTGGCAACCTATC 240  
 QY 1280 CTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339  
 Db 241 CAGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 1340 CTGGAAGTGTGTTGATGGAAGCGTGAAGCTTAACTGGGATGAGACGATACCTATC 1399  
 Db 301 CTGGAAGTGTGTTGATGGAAGCGTGAAGCTTAACTGGGATGAGACGATACCTATC 360  
 QY 1400 ATTCACTCTTGTGTCAGAAATTTGCTGGAAGCTTAAAGTTAACTAACTAACTAACTAA 1459  
 Db 361 ATTCACTCTTGTGTCAGAAATTTGCTGGAAGCTTAAAGTTAACTAACTAACTAACTAA 420  
 QY 1460 AGCGAAGACAGATGGAAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519  
 Db 421 AGCGAAGACAGATGGAAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY	1520	CAGCAGCATGATGCTTAACATGACAAAGACCAAGGTTAACTACTACATCTTTATGAG	1579
DB	481	AAGATATATGTTCCCAACATGACAGGTTCCGAGGGTAAATATATATACATCATATGAG	540
QY	1580	ACTCTGAGAGCATTTCCGGGAAAGACGACCGCAGTCTCGGAGCTTGATATAA	1629
DB	541	ATTCGAGAGGCTCGCAGGAAAGGACGACAGTGTGGAACTTGACAA	590
RESULT 11			
LOCUS	BQ584361		
DEFINITION	EST11859-024-003-K20-SP6 MP1Z-ADIS-024-inflor	linear	EST 06-DEC-2002
ACCESSION	BQ584361		
VERSION	BQ584361.1		
KEYWORDS	EST.		
SOURCE	Beta vulgaris		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Beta.		
REFERENCE	1 (bases 1 to 565)		
AUTHORS	Herwig,R., Schulz,B., Weishaar,B., Hennig,S., Steinfath,M., Drungowski,M., Stahl,D., Wruck,M., Menze,A., O'Brien,J., Lehnach,H. and Radelof,U.		
TITLE	Construction of a 'unigenes' cDNA clone set by oligonucleotide fingerprinting allows access to 25 000 potential sugar beet genes		
JOURNAL	Plant J. 32 (5), 845-857 (2002)		
MEDLINE	22362189		
PUBMED	12472698		
COMMENT	Contact: Weishaar B ADIS DNA core facility at MP1Z Max-Planck-Institute for Plant Breeding Research Carl-von-Linne Weg 10, 50829 Koeln, Germany Fax: 00492215062851 Email: weishaar@mp1z-koeln.mpg.de Insert Length: 565 Std Error: 0.00 Plate: 3 row: K column: 20 Seq primer: SP6; CATACATTGATGACACTATG. Location/Qualifiers 1..565 /organism="Beta vulgaris" /mol_type="mRNA" /contig_id="KMS2320 (double haploid, monogerm breeding line)" /db_xref="GABI:182335" /db_xref="taxon:161934" /clone="024-003-K20" /tissue_type="inflorescence" /lab_host="EMDH108" /clone_lib="MP1Z-ADIS-024-inflor		
FEATURES			
Source			
ORIGIN			
Query Match	21.4%	Score 350.8;	DB 5; Length 565;
Best Local Similarity	77.1%	Pred. NO. 1,2e-93;	
Matches	427;	Conservative 0;	Mismatches 127; Indels 0; Gaps 0;
QY	323	CACGGCTGACAGTGTCTTTACGCCATCACAGATTGGATTCAGGTTACATACAGTCT	382
DB	1	CAAGGCTGGAATGAGTACTCTCTGCAATTAACGAACTTGACCCAGGTTACATACAGTCT	60

OY	383	CTCTTTTCACTGCTCGGAAAGAGGCGCTTAAGTGGTGTGGAGTTGGTATATAGAACAA	442
Db	61	CTCTCTCTCTGTGTGGAAAGATGGGTTAAATGGTGTATATGAGTTTGGTATAGAGCCAA	120
OY	443	ATGCAATTTGTCGCTGTTCCATACGATTTGGAGATTTGTCAACCAACCAAATTTGAAAGACGTG	502
Db	121	ATGCAATTAATGTGTTCTTATGATTTGGAGACTATCACACATCAAAAATTGGAGAAACGAG	180
OY	503	ACCTTTACTTTCACAAAGCTCAAGTTGACCTTTGAAACGTGCTTAAACCTCCGGTGGGGGC	562
Db	181	ACCTTTATTTCCACAAACTCAAGCTGACTTTTGAACATGCTACAAACCTTCTGGGGGGCC	240
OY	563	CTTCTATATGATATTTGGCCCATTCATAGGGTAAATATGTCTTCAGATATCTTTCGAAATGCG	622
Db	241	CGTATATGTGTAGCGCTCATTCACCTTGGAAACCTGTTTTCGATACCTTTCAGAGTGGT	300
OY	623	TGAGGCTTGAATTTGCACCAAAAACATTTATTTGGAAGTGGCTGATACAGATATACCAATGCTT	682
Db	301	TGAAGCTTGAATATGCTCCAAAAGATATATGAATTTGCTGTATTAACATCCATCAGCCCT	360
OY	683	ATTTCGCTGTGGAGCTCTCTCTTCTTGCTTCTGTGAGGCAATCAAACTACTCTCTCTG	742
Db	361	ATTTTGTCTGTGGAGCCCTTTTCTTGTGTGCAAGGAAACATCAAGGCTGCAATTTTGG	420
OY	743	GTTTAAAGTTTGGCGCTTCTCTGTTTCTGAGGGAATCTGCTCGGTTGTTTTCCAATTTCTTTTG	802
Db	421	GTTTCACCTTGGCGGCTTCAGTTTCTGAGGAAACGTCTGCTTAATGTTCAAATTTCTTTTGG	480
OY	803	CGTGGTCAATGTGGCTTAATGCAATTTTCAAAAGAAATTTGCAGGGGTATATAACATCCTGGA	862
Db	481	GTTTCAATCATGTGATGTGGCATTTTCAAAATACCTGTAAACAATGATATAGTACTGGA	540
OY	863	CGCAATTTTCTGGG	876
Db	541	AGCACTTTTCTGAG	554

```

RESULT 12
LOCUS      CD710878
DEFINITION VAA026C02.402285 An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay
CDNA clone VAA026C02 5, mRNA sequence.
ACCESSION  CD710878
VERSION     CD710878.1  GI:32245059
KEYWORDS    EST.
SOURCE      Vitis vinifera
ORGANISM    Vitis vinifera
REFERENCE   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.
REFERENCE   1 (bases 1 to 560)
AUTHORS     Cramer,G.R. and Cushman,J.C.
TITLE       An expressed sequence tag database for abiotic stressed leaves of Vitis vinifera var. Chardonnay
JOURNAL     Unpublished (2002)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jushman@unr.edu
PCR PRIMERS
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 026 row: C column: 02
Seq primer: T7 20mer (forward)
High quality sequence stop: 560.
            Location/Qualifiers
                1..560
                /organism="Vitis vinifera"
                /mol_type="mRNA"
FEATURES
Source

```



```

/db_xref="taxon:29760"
/clone="VVA026C02"
/tissue_type="leaf"
/dev_stage="juvenile and adult"
/clone_1ib="An expressed sequence tag database for abiotic
stressed leaves of Vitis vinifera var. Chardonnay"
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Library construction was performed
according to Stratagene's recommended protocol for the
Lambda UnizapXR vector and cDNA synthesis kit."

```

## ORIGIN

```

Query Match      20.9%; Score 343; DB 6; Length 560;
Best Local Similarity 77.2%; Pred. No. 2.7e-91;
Matches 429; Conservative 0; Mismatches 126; Indels 1; Gaps 1;

```

```

QY 1016 CGAGATGGAATGTGGCTTCCACCCCTTTTGTCTTTCACAGCCCTGTAACATGACGATG 1075
DB 5 CTAAACATGGAATGTGACCTCCACATCAAGTATCTCTTCCCTCTGTGATATACAGATG 64
QY 1076 GGA-CTCTTTCAAGCAATAGAGACTATGACCCAGATAGCAAGAGATGTTACACAG 1134
DB 65 GAACNCCTTTTCAAGCAATAGAGGTTATGATCCAGATAGCAAGAGCTCTTTGACCA 124
QY 1135 TTTAAAGAGTTGATCATGATGATGACCTGTTTAAATCCTCTGATCTCTTGGAGAGACCA 1194
DB 125 TTACAGAAAGTTATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 184
QY 1195 CCTATTAATAAATGTTATTTTGCATATATGTTGCTCATCTTAAGACAGAGTTGTTATAC 1254
DB 185 CTTTAAAGACATTTTGTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
QY 1255 TTTGCCCAAGTGGCAAACTTATCTCTGATTAATTTGATCATGACGATATCATTTACGA 1314
DB 245 TTTGACCAAGTGGCAAACTTATCTCTGATTAATTTGATCATGACGATATCATTTATGAG 304
QY 1315 ACTGAAGTTCCCTGCTGTCAGAGTCTGAGTGTGATGATGATGATGATGATGATGATGATG 1374
DB 305 ATTGAAGATCTTTAGTCTCAAGATCAGAGAAATCTAGTGAAGGAAACCAAGAGCTGCC 364
QY 1375 ACTGGGATGAGACGGTACCCCTATATCTCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTG 1434
DB 365 AGTGGGATGAGACGGTACCCCTATATCTCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 424
QY 1435 AAAGTTAAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1494
DB 425 AAAGTGAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
QY 1495 CTAAATGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1554
DB 485 TTCAATGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
QY 1555 GTTAGTACATTAACCT 1570
DB 545 GTTAGTACATTAACCT 560

```

```

RESULT 13
CA090564          676 bp  mRNA  linear  EST 23-SEP-2003
LOCUS             SCSGAM2105C04.g AM2 Saccharum officinarum cDNA clone SCSGAM2105C04
DEFINITION       5', mRNA sequence.
ACCESSION        CA090564
VERSION          CA090564
KEYWORDS         EST
SOURCE           Saccharum officinarum
ORGANISM         Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE        1 (bases 1 to 676)
AUTHORS          Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

```

## TITLE

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parud@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 105 row: C column: 04

Seq primer: 77 Promoter Primer.

Location/Qualifiers

1. 676

/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCSGAM2105C04"

/lab\_host="DH10B"

/clone\_1ib="AM2"

/note="Organ: Apical meristem and tissues surrounding of

immature plants; Vector: pSPori; Site\_1: SalI; Site\_2:

NotI; An unidirectional cDNA library generated from

[Apical meristem and tissues surrounding of immature

plants]. cDNA was prepared from poly(A<sup>+</sup> mRNA using

SuperScript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a separose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at <http://sucest.lad.ic.unicamp.br/public>"

## ORIGIN

```

Query Match      20.2%; Score 331.8; DB 6; Length 676;
Best Local Similarity 75.5%; Pred. No. 6.9e-88;
Matches 425; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

```

```

QY 79 GCGGTGGAGATGAGACCGAGTTTACCGGCACTACGAGCTATCGGGATATATCTT 138
DB 114 GCGGCGGGGCGGAGCGGCGCGGAGTTTCAATTCAGGAAGCTTCGGGATCATATC 173
QY 139 CCGGATTTGGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 198
DB 174 CCGGCGTTCCGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 233
QY 199 TTGACTTCAATCCGCTGACCTGATGCTAGACACCACTAAGCTTCTTCTGCTGCT 258
DB 234 TTGACTTCAATCCGCTGACCTGATGCTAGACACCACTAAGCTTCTTCTGCTGCTA 293
QY 259 AACTGCTGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318
DB 294 AATTGCTGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 353
QY 319 AAGTCACGCGCTGACAGTGTCTTTCACGACCATCAAGATTTGATCCAGTTTACATACA 378
DB 354 AAGTCACGCGCTGACAGTGTCTTTCACGACCATCAAGATTTGATCCAGTTTACATACA 413
QY 379 GGTCTCTTTTCTACTGCTGGAAGAGTGGCTTAAAGTGTGTGTTGAGTTTGTATAGAA 438
DB 414 GGTCTCTCTCTTCTACTGCTGGAAGAGTGGCTTAAAGTGTGTGTTGAGTTTGTATAGAA 473
QY 439 GCAATGCAATTTGCGCTGCTCAATAGATTTGGAATGTCACCAACCAATTTGGAAGAG 498
DB 474 GCTAATGCAATTTGCGCTGCTCAATAGATTTGGAATGTCACCAACCAATTTGGAAGAG 533
QY 499 CGTGAACCTTACTTTCACAGACTGAGTTGACCTTTGAACTGTTTAAACT-CCGTGG 557
DB 534 CGAGATCTGCTTTTTCACAGACTGAGTTGACCTTTGAACTGTTTAAACTCCCGGGG 593
QY 558 CGGCGCTTCTATATGATTTTGGCCATTCATGAGGATATATGTTTTCAGATCTTCTGGA 617

```



Db 594 AGGGCCATCTTGTGATTTGCTCTTCGATGCAATATATGTTCGCTACTTTTGA 653  
Qy 618 ATGGCTAGGCTAGAAATTCAC 640  
Db 654 ATGCTTGAAACTGAAATATGCC 676

RESULT 14  
AI974291 534 bp mRNA linear EST 26-AUG-1999  
LOCUS AI974291  
DEFINITION T1102936 KV0 Medicago truncatula cDNA clone pKV0-1B14, mRNA  
sequence.  
ACCESSION AI974291  
VERSION AI974291.1 GI:5776732  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.  
1 (bases 1 to 534)  
Vandenbosch, K., Hur, J., Moore, J., Beremand, P., Peng, H. and Ellis, L.  
ESTs from uninoculated roots of Medicago truncatula (1999b)  
Unpublished (1999)  
Contact: Vandenbosch K  
Department of Plant Biology  
University of Minnesota  
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
Tel: 612 624 2755  
Fax: 612 625 1738  
Email: kvandenbosch@umn.edu  
Other name: 07-B-7; date: 6/30/99; Submitted to the Database of  
Expressed Sequence Tags (dbEST) on 08/25/99; More information is  
available at "http://chryslr.tamu.edu/medicago".  
Seq primer: SKmod (CTA GAA CTA GTG GAT CC).  
Location/Qualifiers  
1..534  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pKV0-1B14"  
/tissue\_type="Seedling roots"  
/dev\_stage="Immediately prior to inoculation with  
Sinothizobium meliloti (0 hour)"  
/lab\_host="E.coli strain XL0LR"  
/clone\_1lb="KV0"  
/note="Vector: Bluescript SK -, Site 1: EcoRI, Site 2:  
XhoI. cDNA was prepared from polyA+ enriched RNA. The cDNA  
was directionally ligated into the Unizap XR vector from  
Stratagene and packaged using Gigapack III Gold packaging  
extracts. Plasmids containing cDNA inserts were excised  
from the recombinant lambda-Zap phage using Ex-assist  
helper phage and propagated in XL0LR cells."

ORIGIN  
Query Match 19.3%; Score 316.4; DB 1; Length 534;  
Best Local Similarity 76.3%; Pred. No. 2.7e-83;  
Matches 389; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 989 TTACGAAACAGCTCTAGTCAATGACAGACATGATGAGCCCTTCCACCTTTCT 1048  
Db 25 TTTCGAATATCCCGAGGCCAATCTGTCCGAAATGAAATGACCTACTCAATAT 84  
Qy 1049 CTTTACAGCCCGGTGAATGACATGAGTCTTTTCAAGCAATAGAGCTATGACC 1108  
Db 85 CTTTCTACACTCTGTAAATAGCGAGTGGCTCTTTTCAAGCAATGAAATATGACC 144  
Qy 1109 CAGATAGCAAGAGATGTTACACCACTTAAAGAGTTGATATGATGACCTCTTTT 1168  
Db 145 CAGACAGCAAGAGACTCTTGTACCAATTAGAGAAATCATATCTGTGTATCTCTGATCTTA 204

Qy 1169 ATCTCTGATCTCTTGGAGAGACCACTATATATATTTTGCATATATATGCTC 1228  
Db 205 ATCCACTTACACTTGGATTCGCCACCAATATATATATATATATATATGACAGTA 264  
Qy 1229 ATCTAAGACAGAGTTGTTATTAATTTTGGCCCAAGTGGCAACCTTATCTGATAT 1288  
Db 265 ATTCAAAGCAAGAGTTGGTTACTCTTGGCCCTAGCGGCAAGCCGATCTCTGATMACT 324  
Qy 1289 GGATCATCAGGATATCATTTTACGAACTGAAGGTTCCTGCTGTCATGAGTGTGAAC 1348  
Db 325 GGATCATTCAGATATCTGTTATGATGACGAAAGATCTTATCATCAAGGTACGAAATC 384  
Qy 1349 TGGTATGAGGAGCCCTGACCTATTAATGAGGAGAGCGGTACCTATCACTTACTCT 1408  
Db 385 TGGTATGAGGAGCCCTGATCTTAATGAGGAGATGAGCGGTTCATACACTCTTAT 444  
Qy 1409 CTTGTTGCAAGATTTGGCTCGACCTTAAGTTAATTAATGAGCTCTCCCAAGCAAG 1468  
Db 445 CTTGTTGCAAAATTTGGCTCGACCAAGTTGATTAAGATTAAGAGCTCTCTCAGTCAAG 504  
Qy 1469 ACGATGAAAGCAGCTACATGAGACTTA 1498  
Db 505 ACGATGCTCAGATGTACAGATTGATTGA 534

RESULT 15  
BI267124 628 bp mRNA linear EST 18-JUL-2001  
LOCUS BI267124  
DEFINITION NF097C07IN1F1054 Insect herbivory Medicago truncatula cDNA clone  
NF097C07IN 5', mRNA sequence.  
ACCESSION BI267124  
VERSION BI267124.1 GI:14871872  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.  
1 (bases 1 to 628)  
Korth, K., Scott, A.D., Harris, A.R., Gonzalez, R.A., Bell, C.J.,  
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula insect herbivory library  
Unpublished (2000)  
Contact: Korth K  
Dept. of Plant Pathology  
University of Arkansas  
217 Plant Science Building, Fayetteville, AR 72701, USA  
Tel: 501 575 5191  
Fax: 501 575 7601  
Email: Kkorth@comp.uark.edu  
Insert Length: 628 Std Error: 0.00  
Plate: 097 row: C column: 07  
Seq primer: TCACACAGAAACAGCTATGAC.  
Location/Qualifiers  
1..628  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF097C07IN"  
/tissue\_type="local and systemic leaves"  
/dev\_stage="mature"  
/clone\_1lb="Insect herbivory"  
/note="Vector: Lambda Zap; Library was produced from fully  
expanded M. truncatula leaves of plants fed upon by  
Spodoptera exigua (beet armyworm) for 24 hours. Systemic  
(undamaged leaves from injured plants) and wounded leaves  
were harvested and pooled."

ORIGIN  
Query Match 18.7%; Score 306.6; DB 4; Length 628;  
Best Local Similarity 79.2%; Pred. No. 2.6e-80;

Matches 374; Conservative 0; Mismatches 97; Indels 1; Gaps 1;

OY	108	CGACTACTGAGAGTATCCGGATATATCATTCGGGATTTGGTCGACGACGTAAGAGC	167
Db	149	CGATTACTCTAAACTCTCCGGCATTATATCCAGGCTTGGCTTACTCAGCTTCGAGC	208
OY	168	GTGGTCGATTCCTTGACTGTCATACACTCCGTTGACCTCAATCCGCTCGACCTCGTATG	227
Db	209	ATGGTCTATCTTGACTGCTCCCTTACTCTCCGCTAGATTCAATCCCTCGATTTGGTTTG	268
OY	228	GCTAGACACCACTAAGCTTCTTTGCTGTCAACTGCTGTTTAAGTATGATGCTAGTA	287
Db	269	GCTCGACACCAACCAACTTCTTTCTGCTGTGAATTTGGCTTAAGTCATGTTGGA	328
OY	288	TCCCTATATCAAAACAGACCAATCCGAGTGTAAAGTCACGCGCTGACAGTGTCTTTACAC	347
Db	329	CCCTTACAATCAAAACAGATCATCTGATTGTAAATCCGCTCGATAGCGGTCTTTCCGG	388
OY	348	CATCAGAGATTGATCCAGGTTACATACAGGTCTCTTTCTACTGTCTGGAAGAGTG	407
Db	389	AATTACAGAGCTTGATCCAGGTTATATACAGAGACCTCTTCGTGTATGGAAAGAAATG	448
OY	408	GCTTAAGTGTGTGTGATTGTTGTTATAGAGCAAAATGCAATTTGCTGTTCATACGA	467
Db	449	GATTAAAGTGTGTATTTGAATTTGGCATANAAGCTAATGCAATATTTGCTTCTTATGA	508
OY	468	TTGAGATTGTACCAACCAAAATGGAGAGCGTGAACCTTTAATTTCACAGCTCAAGTT	527
Db	509	TTGGAGACTGTACCATCTGCTTGAGAGACGAGACCTTACTTCAATAGCTTAATTT	568
OY	528	GACCTTTGAAAGCTTTTAAATCCGTGGCGGCCCTTCTATAGTATTTGCC	579
Db	569	GACATTGTGAGCTGCTTT-CAACTTCGAGAGNGGCCCTCTTAAATTTTGGC	619

Search completed: November 9, 2005, 08:31:22  
Job time : 5649 secs